



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

Mesure de la fragilité et détection de chutes pour le maintien à domicile des personnes âgées

THÈSE

présentée et soutenue publiquement le 15 septembre 2014

pour l'obtention du

Doctorat de l'Université de Lorraine
(mention informatique)

par

Amandine Dubois

Composition du jury

<i>Président :</i>	Jean Paysant	Professeur à la Faculté de Médecine de Nancy
<i>Rapporteurs :</i>	Philippe Lalanda Jean-Louis Vercher	Professeur à l'Université Joseph Fourier de Grenoble Directeur de Recherche au CNRS à Marseille
<i>Examineurs :</i>	Jacques Demongeot Jacques Duchêne Michel Vacher	Professeur à l'Université Joseph Fourier de Grenoble Professeur à l'Université de Technologie de Troyes Ingénieur de recherche au CNRS de Grenoble
<i>Directeur de thèse :</i>	François Charpillet	Directeur de recherche INRIA

Mis en page avec la classe thesul.

Remerciements

Je tiens à remercier tout d'abord François Charpillat de m'avoir donné l'opportunité de faire une thèse en Informatique et de m'avoir accordé sa confiance.

Je remercie également les membres de mon jury : monsieur Jean Paysant pour avoir présidé ce jury, messieurs Philippe Lalanda et Jean-Louis Vercher de m'avoir fait l'honneur de rapporter ce travail de thèse et messieurs Jacques Demongeot, Jacques Duchêne et Michel Vacher pour leurs remarques et leurs questions constructives.

Merci également à Abdallah, Manel, Marion, Olivier et à tous ceux que j'ai pu côtoyer pendant ces trois ans, collègues de bureau, de couloir, de bâtiment, pour les discussions, leurs suggestions et surtout pour leur bonne humeur.

Je remercie mes parents et mon frère qui m'ont toujours soutenue et encouragée et sans lesquels je ne serais pas parvenue à accomplir ce travail de thèse.

Je remercie mes grands-parents pour avoir été les premiers « cobayes » et surtout pour m'avoir donné l'envie de travailler sur ce sujet.

Je te remercie Cédric pour ton soutien permanent.

Sommaire

Table des figures	xi
-------------------	----

Liste des tableaux	xv
--------------------	----

Introduction	1
1 Contexte	3
2 Objectifs	4
3 Approche proposée	4
3.1 Problématique	4
3.2 Contributions	5

Partie I Approches pour la mesure de la fragilité et la détection des chutes	9
---	----------

Introduction	11
---------------------	-----------

Chapitre 1

Panorama des approches pour l'analyse de la fragilité à l'hôpital

1.1 L'œil expert à travers les tests cliniques	13
1.2 Outils d'analyse du mouvement	14
1.2.1 Systèmes avec capteurs embarqués sur la personne	14
1.2.2 Systèmes sans capteur embarqué sur la personne	16
1.3 Conclusion	16

Chapitre 2

Panorama des systèmes de détection de chutes au domicile

2.1 Systèmes avec capteurs embarqués sur la personne	19
--	----

2.1.1	Détecteurs de chutes automatiques	19
2.1.2	Systèmes d'alerte manuels	21
2.2	Systèmes ambiants	21
2.2.1	Capteurs environnementaux : la télé-assistance	21
2.2.2	Systèmes à base de caméra : la vidéo-vigilance	22
2.2.3	Capteurs au sol	22
2.3	Conclusion	22
Conclusion		25

Partie II Algorithmes pour l'évaluation et la sécurisation au domicile 27

Introduction	29
---------------------	-----------

Chapitre 3

Détecter et suivre une personne avec une caméra : état de l'art

3.1	Détection de la personne	33
3.1.1	Détecteur de points d'intérêt	34
3.1.2	Soustraction du fond	34
3.1.3	Segmentation	35
3.1.4	Apprentissage supervisé	35
3.2	Suivi de la personne	35
3.2.1	Modèles d'objets animés	35
	Différentes représentations	35
	Caractéristiques pour obtenir la représentation	37
3.2.2	Suivi de l'objet	37
	Objet représenté par des points	38
	Objet représenté par des formes géométriques	39
	Objet représenté par sa silhouette	39
3.3	Caméra de profondeur	40
3.4	Conclusion	41

Chapitre 4

Application à la détection et au suivi de personnes avec une caméra RGB-D

4.1	Détecter la personne	43
-----	--------------------------------	----

4.1.1	Extraction et apprentissage du fond	44
4.1.2	Extraction du sol et changement de repère	44
	Etape 1 : Redressement de la scène	44
	Etape 2 : Sélection des points bas	45
	Etape 3 : Calcul de l'équation du sol	45
	Etape 4 : Construction du repère sol	46
	Projection d'un point dans le repère sol	46
4.1.3	Extraction des points mobiles	46
4.1.4	Extraction des objets mobiles	47
4.1.5	Détection de la personne	49
4.2	Suivre la personne	49
4.2.1	Représentation de la personne	49
4.2.2	Suivi du centre de masse	51
	Calcul du centre de masse	51
	Description de la trajectoire du centre de masse	51
	Lissage de la trajectoire du centre de masse	52
	Paramètres dynamiques	53
4.2.3	Suivi de la distribution verticale	53
4.3	Conclusion	55

Chapitre 5

Mesurer à domicile les paramètres de la marche

5.1	Définition et description de la marche	58
5.1.1	Définition de la marche	58
5.1.2	Description du cycle de la marche	58
	Phase d'appui	59
	Phase d'oscillation	59
5.2	Indicateurs de l'analyse de la marche	59
5.2.1	Paramètres spatio-temporels	59
	Principaux paramètres spatiaux	59
	Principaux paramètres temporels	60
5.2.2	Paramètres cinématiques	60
5.2.3	Paramètres dynamiques	61
	Les forces de réaction du sol	61
	Pressions plantaires	62
5.3	Sélection des indicateurs	62
5.3.1	Troubles de la marche et risques de chute	64

5.3.2	Troubles de la marche et risques d'hospitalisation	64
5.3.3	Troubles de la marche et troubles cognitifs	64
5.4	Extraction des indicateurs	65
5.4.1	Longueurs de pas	65
5.4.2	Cadence	66
5.4.3	Vitesse de marche	66
5.4.4	Résultat	66
5.5	Conclusion	67

Chapitre 6

Reconnaitre l'activité humaine en temps réel par l'analyse d'images de profondeur

6.1	Approches pour détecter l'activité	70
6.1.1	Approches non paramétriques	70
6.1.2	Approches volumétriques	70
6.1.3	Approches paramétriques	71
6.2	Positionnement	71
6.3	Proposition d'un modèle de l'activité	73
6.3.1	Modèle avec fonction d'observation seule	74
	Inférence	74
6.3.2	Utilisation d'un MMC unique	75
	MMC à 8 états et 3 observations	75
	MMC à 9 états et 5 observations	76
	Inférence	79
6.3.3	Utilisation de MMC combinés	79
	Système de 8 MMC	79
	Système de 9 MMC	79
	Inférence	80
6.4	Conclusion	81

Conclusion	83
-------------------	-----------

Partie III Evaluation du système	85
---	-----------

Introduction	87
---------------------	-----------

Chapitre 7

Expérimentations pour l'analyse de la marche

7.1	Première évaluation du système	89
7.1.1	Description de l'expérimentation	89
7.1.2	Résultats	91
	Angle de la caméra	91
	Choix du filtre	93
7.1.3	Conclusion	94
7.2	Évaluation globale du système	94
7.2.1	Description de l'expérimentation	94
7.2.2	Résultats	95
	Résultats basés sur l'erreur de chaque séquence	96
	Résultats basés sur la moyenne de chaque paramètre de la marche	96
7.3	Discussion et conclusion	97

Chapitre 8

Expérimentations pour la reconnaissance d'activité

8.1	Cas de la personne entièrement visible	99
8.1.1	Description de l'expérimentation	99
	Protocole expérimental	99
	Base d'apprentissage et base de test	100
8.1.2	Résultats	102
	Représentation des bonnes classifications et des fausses détections	102
	Comparaison entre les algorithmes <i>Forward-Backward</i> et de Viterbi	103
	Modèle avec fonction d'observation seule	105
	MMC à 8 états et 3 observations	105
	MMC à 9 états et 5 observations	107
	Système de 8 MMC à 2 états et 3 observations	107
	Système de 8 MMC à 3 états et 3 observations	107
	Système de 9 MMC à 2 états et 5 observations	107
	Système de 9 MMC à 3 états et 5 observations	108
	Résultats pour l'activité debout	108
8.1.3	Conclusion	108
8.2	Cas de la personne partiellement visible	109
8.2.1	Description de l'expérimentation	109
8.2.2	Résultats	109
8.2.3	Conclusion	113

Chapitre 9

Tests en situations réelles

9.1	Test de l'algorithme d'analyse de la marche	115
9.2	Test de l'algorithme de détection d'activité	116
9.3	Conclusion	118

Conclusion	121
-------------------	------------

Partie IV Système pour le maintien à domicile 123

Introduction	125
---------------------	------------

Chapitre 10

Vers une implantation au domicile

10.1	Couplage des fonctionnalités actuelles du système	127
10.2	Cartographie des habitudes dans l'environnement	129
10.2.1	Limitations du système et contraintes de l'environnement	129
10.2.2	Détection des cas d'occlusions partielles de la personne	130
10.2.3	Cartographie des habitudes	133
10.2.4	Détection des anomalies	134
10.3	Conclusion	135

Chapitre 11

Autres perspectives d'utilisation du système

11.1	Différenciation des personnes	137
11.2	Couplage de plusieurs capteurs	138
11.3	Observance des patients	139
11.4	Prévention de certains troubles	139
11.5	Conclusion	140

Conclusion	141
-------------------	------------

Conclusion générale et perspectives	143
Annexe	149
Annexe A Modèle de Markov caché	149
A.1 Définition des modèles de Markov cachés	149
A.1.1 Processus de Markov	149
A.1.2 Modèle de Markov caché	150
A.2 Algorithmes d'apprentissage et d'inférence	150
A.2.1 Inférence pour le cas du MMC unique	150
Algorithme <i>Forward-Backward</i>	150
Algorithme de Viterbi	153
A.2.2 Apprentissage et inférence pour les systèmes de plusieurs MMC	155
Apprentissage	155
Inférence	156
Publications	157
Références	159

Table des figures

1	Pyramide des âges en 2007 et 2060 selon une étude INSEE de 2010.	3
1.1	Représentation de la marche humaine en diagramme bâton par Marey.	14
1.2	Le système Walkmeter© est inspiré du locomètre de Bessou.	15
1.3	Le tapis roulant instrumenté de la société Biometrics.	17
1.4	Le tapis actimétrique de la marque GaitRite.	17
2.1	Système d’alerte de Vitalbase.	20
2.2	Système de détection de chute de la société Vigilio.	20
2.3	Système d’analyse du mouvement par le sol d’INRIA à Nancy.	23
3	Vision pour la conception d’un système permettant le maintien à domicile des personnes âgées.	30
3.1	Détection des points d’intérêts dans une scène.	34
3.2	Différentes représentations de l’objet à suivre : (a) silhouette, (b) contour, (c) points, (d) centre géométrique, (e) ellipse, (f) rectangle, (g) plusieurs ellipses et (h) squelette.	36
3.3	Détection des contours dans une scène.	37
3.4	Différentes hypothèses utilisées pour le suivi d’un objet : (a) proximité, (b) vitesse maximale, (c) mouvement similaire, (d) constance de la vitesse et de la direction, (e) rigidité de l’objet.	38
3.5	Transformation paramétrique d’une ellipse.	39
3.6	Modèle d’apparence.	40
3.7	Le capteur Kinect de Microsoft.	40
4.1	Étapes pour détecter la personne du fond et la suivre dans le temps.	44
4.2	L’extraction du sol est représentée en rouge.	47
4.3	a) Les cases bleues représentent des points mobiles et les cases vertes les points appartenant au fond, b) le filtre érosion est appliqué, c) le filtre dilatation est appliqué.	48
4.4	a) Les cases colorées représentent des points mobiles, b) l’algorithme assigne un nombre aux points mobiles, c) les points avec un nombre plus grand que ses voisins sont modifiés, d) le résultat final est atteint.	48
4.5	Déplacement horizontal théorique du centre de masse.	50
4.6	Déplacement vertical théorique du centre de masse.	50

4.7	Déplacement du centre de masse extrait à partir de notre algorithme pour une personne marchant en ligne droite perpendiculairement à la caméra. La montée et la descente en début et fin de trajectoire correspondent aux entrées et sorties de la personne dans le champ de vision de la caméra.	51
4.8	Comparaison entre la trajectoire du centre de masse et le nombre de pixels mobiles appartenant à une personne.	52
4.9	a) Courbe non lissée, b) courbe lissée avec le filtre passe-bas, c) courbe lissée avec le filtre de Kalman.	54
4.10	Trajectoire de l'écart type vertical pour une personne qui marche.	54
4.11	Comparaison entre la trajectoire du point maximum filtrée et non filtrée pour une personne qui marche.	55
5.1	Représentation du cycle de la marche selon Viel (2000).	58
5.2	Représentation des paramètres spatiaux.	60
5.3	Représentation des paramètres cinématiques.	61
5.4	Représentation des forces exercées lorsqu'une personne marche.	62
5.5	Force de réaction du sol selon l'axe vertical (bleu), antéro-postérieur (vert) et média-latéral (rouge) lors d'un cycle de marche selon Kirtley.	63
5.6	Trajectoire des pressions plantaires lors de la marche.	63
5.7	Correspondance entre le cycle de la marche et la trajectoire du centre de masse sur l'axe vertical.	66
5.8	Extraction des paramètres de la marche à partir du centre de masse.	67
6.1	Positionnement du problème de la reconnaissance d'activité au domicile.	71
6.2	Représentation graphique d'une chaîne de Markov cachée, avec X_t le processus Markovien caché et O_t la séquence d'observations correspondantes.	73
6.3	Modèle à 8 états avec fonction d'observation seule.	75
6.4	MMC à 8 états avec des transitions définies manuellement.	77
6.5	MMC à 9 états avec des transitions définies manuellement.	78
6.6	Système de 8 MMC à deux états chacun.	80
6.7	Système de 9 MMC à deux états chacun.	81
7.1	Méthode utilisée comme référence pour mesurer les longueurs de pas.	90
7.2	Positionnement des caméras : la plus basse avec un angle de 0° et la plus haute avec un angle de -27°	91
7.3	Représentation des erreurs sur les longueurs de pas obtenues à partir d'une caméra avec un angle de -27° et avec un angle de 0°	92
7.4	Représentation des erreurs sur les longueurs de pas extraites à partir du filtre de Kalman et du filtre passe-bas.	94
7.5	Expérience comprenant un tapis actimétrique.	95
7.6	En phase de double appui, de nombreux points mobiles sont détectés dans la partie basse de la silhouette quand le sujet porte une jupe.	98
7.7	Comparaison des oscillations de la trajectoire du centre de masse.	98
8.1	Plan de l'appartement expérimental avec le positionnement de la caméra.	100
8.2	Suivi des situations, effectuées par un sujet, à différents pas de temps.	101
8.3	Différence dans l'analyse faite par les algorithmes <i>Forward-Backward</i> et de Viterbi, pour une séquence où la personne s'assoit puis monte sur une chaise.	104
8.4	Comparaison d'une séquence où la personne s'assoit analysée par deux modèles.	105

8.5	Sujet s'allongeant différemment de ce qui a été appris dans la base de donnée. . .	106
8.6	Sujet s'asseyant différemment de ce qui a été appris dans la base de donnée. . . .	106
8.7	Suivi des situations où le sujet n'est pas entièrement visible à différents pas de temps.	111
8.8	Résultat de l'analyse avec le modèle à 9 états et 5 observations, pour une personne réalisant des demi-tours.	112
9.1	Utilisation de notre système avec une patiente, lors de sa consultation à l'hôpital.	116
9.2	Représentation du centre de masse, sur l'axe vertical, de quatre personnes âgées.	117
9.3	Détection de l'activité « Marche », pour une personne de 81 ans marchant à son domicile.	117
9.4	Représentation du centre de masse, sur l'axe vertical, pour différentes activités réalisées par une personne âgée de 81 ans.	118
10.1	Les différentes actions réalisées par l'algorithme selon l'état dans lequel se trouve la personne d'après le MMC.	128
10.2	Personne sortant du champ de vision et étant à moitié détectée.	129
10.3	Courbes des vraisemblances pour différentes activités réalisées par un sujet entièrement et partiellement visible.	132
10.4	Séquence dans laquelle le sujet réalise plusieurs activités.	133
10.5	Carte de l'environnement représentant les habitudes apprises d'une personne (en blanc : zones non explorées, en bleu : zones passives, en vert : zones actives et en rouge : zones se trouvant derrière un meuble).	134
10.6	Carte de l'environnement sur laquelle est représentée la dernière position du centre de masse perçue par le système.	135
11.1	Proposition d'emplacement de plusieurs caméras permettant de couvrir tout un appartement.	138
11.2	Représentation du centre de masse, sur l'axe vertical, d'une personne de 90 ans ayant des problèmes d'équilibre	139
A.1	Exemple de modèle de Markov à deux états.	149
A.2	Représentation graphique d'une chaîne de Markov.	150
A.3	Représentation graphique d'une chaîne de Markov cachée, avec X_t le processus Markovien caché et O_t la séquence d'observations correspondantes.	150
A.4	Représentation du calcul $\alpha_{t+1}(j = 1)$ par l'algorithme <i>Forward</i> pour un modèle à trois états.	152
A.5	Représentation du calcul pour l'état $\beta_t(i = 1)$ par l'algorithme <i>Backward</i> pour un modèle à trois états.	153
A.6	Représentation du calcul $\delta_t(j = 1)$ et $\psi_t(j = 1)$ par l'algorithme de Viterbi pour un modèle à trois états.	154
A.7	Représentation du calcul de l'algorithme de Viterbi avec la rétro-propagation (en vert) pour quatre pas de temps successifs et trois états.	154

Liste des tableaux

7.1	Comparaison des résultats, entre le tapis et la caméra, en fonction de l'angle de la caméra.	92
7.2	Précision des longueurs de pas fournies par l'algorithme, obtenues avec un filtre de Kalman et avec un filtre passe-bas (pour tous les pas, pour les pas « normaux » et pour les petits pas).	93
7.3	Erreurs moyennes des longueurs (Long), cadences (Cad) et vitesses (Vit) pour chaque situation.	96
7.4	Comparaison des paramètres de la marche obtenus avec le tapis et avec notre algorithme utilisant la caméra.	97
8.1	Classification des activités par l'algorithme selon le modèle utilisé.	103
8.2	Fausse détection effectuée par l'algorithme selon le modèle utilisé.	104
8.3	Classification de l'activité « Debout » selon le modèle utilisé.	108
8.4	Détection d'activité quand la personne n'est pas entièrement visible pour différents modèles.	113

Introduction

1 Contexte

L'un des sujets majeurs des prochaines années dans les pays développés sera le vieillissement de la population. En France, l'INSEE [Blanpain and Chardon, 2010] prévoit que le nombre de personnes de plus de 60 ans augmentera de 10,4 millions entre 2007 et 2060 pour atteindre le nombre de 23,6 millions de personnes en 2060, soit une hausse de 80 % en 53 ans. Le nombre de personnes de plus de 75 ans passera de 5,2 millions en 2007 à 11,9 millions en 2060. Quant aux personnes de plus de 85 ans, leur nombre passera de 1,3 à 5,4 millions. Ces chiffres sont représentés sur la pyramide des âges fournie par l'INSEE à la figure 1.

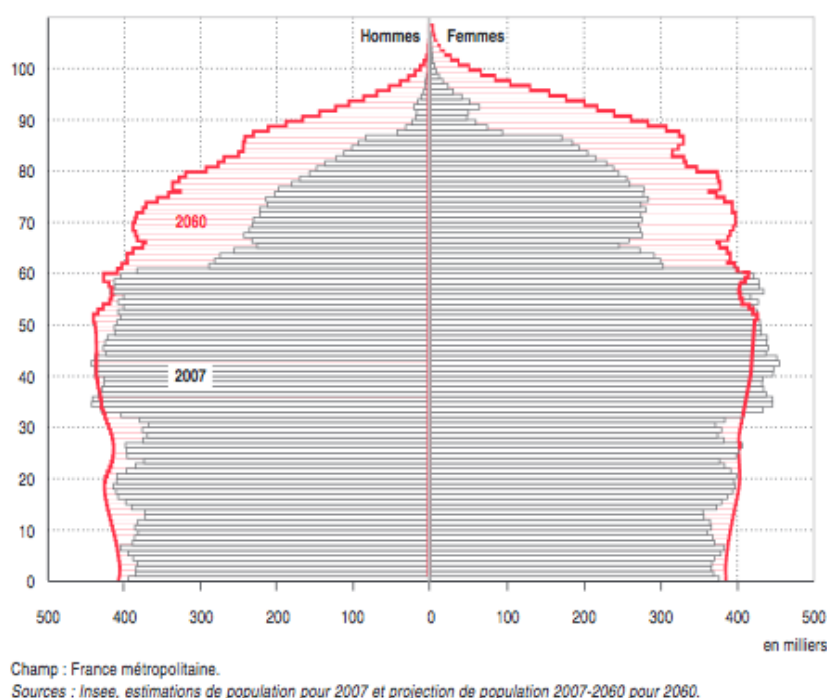


FIGURE 1 – Pyramide des âges en 2007 et 2060 selon une étude INSEE de 2010.

Une des conséquences de l'augmentation du vieillissement de la population est l'augmentation du nombre de personnes âgées dépendantes¹. Selon l'INSEE en 2004, 1 200 000 personnes seraient dépendantes en 2040 contre 800 000 aujourd'hui. La problématique qui découle de ces chiffres est la prise en charge de ces personnes. En effet, le nombre d'instituts spécialisés ne devrait pas permettre de toutes les accueillir. Actuellement, les personnes ont tendance à rester au domicile. 92 % des plus de 75 ans vivent chez eux et 75 % pour les plus de 85 ans, soit pour des raisons liées au coût qu'implique un placement, ou soit parce que la personne elle-même préfère rester chez elle (80 % des personnes âgées souhaitent vivre le plus longtemps possible au domicile [Webe, 2006]).

Le vieillissement de la population est donc un problème majeur et, partant de ce constat, cette thèse cherche à apporter une contribution au maintien à domicile des personnes âgées par des dispositifs technologiques.

1. La dépendance s'évalue en fonction de la capacité de la personne à accomplir les actes de la vie quotidienne. L'outil Aggir est une grille nationale évaluant le degré de dépendance chez les personnes âgées à partir des activités qu'elles peuvent effectuer seules comme s'habiller, se laver, etc.

2 Objectifs

Notre objectif est de développer un dispositif simple et facile à installer, permettant de favoriser le maintien à domicile des personnes âgées. Il s'agit de proposer une démarche de prévention de l'aggravation de la fragilité de la personne et d'aider à la sécuriser chez elle.

Pour savoir si la personne est en sécurité à son domicile, il est nécessaire de connaître son niveau d'autonomie, de savoir si la personne est fragile, si elle peut se débrouiller dans son logement (se faire à manger, se laver, etc.). Pour cela, l'idée est de développer un outil d'évaluation permettant de mesurer le risque initial et l'évolution du risque que prend la personne en restant chez elle. Ces mesures permettront alors de prendre des actions de prévention comme diriger la personne vers un médecin spécialisé, modifier son environnement pour le rendre plus sûr, adapter son traitement, etc. L'évaluation du degré d'autonomie peut se faire au quotidien en installant un système à domicile. Cette évaluation constitue une information supplémentaire à l'analyse faite par le médecin lors des consultations, car l'information recueillie au domicile constitue une mesure en milieu écologique forcément différente de l'évaluation qui peut être faite en milieu hospitalier.

La sécurisation de la personne passe également par la mise en place de systèmes d'alerte, car maintenir les personnes âgées à domicile n'est pas sans risque. En effet, 61 % des accidents surviennent au domicile et 85 % de ces accidents sont dus aux chutes. La chute est la première cause de mortalité des plus de 65 ans provoquant chaque année 10 000 décès [Alexander *et al.*, 1997]. Elle peut causer des blessures physiques comme des fractures (en particulier du col du fémur), des luxations, des plaies, etc. Les blessures physiques peuvent entraîner une hospitalisation prolongée, voire même le placement en institut (40 % des sujets âgés hospitalisés pour chute sont ensuite orientés vers une institution). Une cascade de complications psychologiques peut également survenir suite à la chute elle-même et/ou du fait d'être resté immobilisé au sol en état d'impuissance sans pouvoir se relever. En effet, près de 50 % des personnes ayant chuté ont besoin d'une aide pour se relever [Tinetti *et al.*, 1993]. L'anxiété d'une deuxième chute peut entraîner chez les personnes une tendance à réduire leur mobilité et à minimiser leur activité [Gaxatte *et al.*, 2011] (en limitant leur sortie et en restant plus souvent assise ou allongée). Cette inactivité provoque une perte d'autonomie. Des études ont montré que la perte d'autonomie pouvait directement être corrélée au temps passé au sol, 1h au sol provoque 10 % de perte d'autonomie, 3h, 30 %, 6h, 50 % et 12h, 70 %. La chute est un marqueur de fragilité et ses conséquences physiques et surtout psychologiques aggravent l'état de fragilité initiale de la personne. Cet engrenage est confirmé par des chiffres montrant que le nombre de récidives est très important : après une première chute le risque de retomber à nouveau est multiplié par vingt. Il est donc nécessaire, pour sécuriser la personne, d'être en mesure de détecter les chutes. Un tel détecteur pourrait permettre, d'une part, de limiter les conséquences physiques et psychologiques de la chute et, d'autre part, de mettre la personne en confiance pour ainsi ralentir les effets de l'inactivité (réduire l'effet engendré par le cercle vicieux). La détection précoce est cruciale, le fait d'obtenir une aide rapidement après une chute diminue le risque d'hospitalisation de 26 % et la mort de plus de 80 %.

3 Approche proposée

3.1 Problématique

Dans ce travail de thèse, nous nous intéressons à développer un système couplant l'évaluation de la fragilité et la sécurisation de la personne à son domicile.

Sécuriser la personne à son domicile implique de pouvoir détecter les éventuelles chutes, ou

comportements pouvant mener à une chute (comme monter sur une chaise), pour ainsi alerter des personnes extérieures et permettre d’agir au plus vite sur les conséquences physiques et psychologiques. La prise en charge rapide d’une personne ayant chuté limite l’aggravation de la perte d’autonomie.

L’évaluation de la fragilité de la personne permet de mesurer sa capacité à être autonome à son domicile. Une personne est fragile, selon la définition de Fried *et al.* [Fried *et al.*, 2001], si apparaît trois critères ou plus parmi les cinq suivants : une perte de poids involontaire, une faiblesse dans la force de préhension, une faible endurance, une vitesse de marche lente et une faible activité physique. La vitesse de marche diminuée et la faible activité sont les deux critères auxquels nous nous sommes intéressés pour l’évaluation de la fragilité. Notre approche vise à analyser les indicateurs de marche pour ainsi mesurer leur évolution au cours du temps et analyser l’activité de la personne pour pouvoir comptabiliser le temps passé à réaliser différentes activités. Par exemple, la perte d’autonomie, analysée à travers l’évaluation de l’activité de la personne, peut s’exprimer en fonction du temps passé debout comparé au temps passé assis et allongé sur un canapé ou un lit. Les changements d’habitudes sont également des indicateurs importants de l’évolution de la fragilité.

Nous nous sommes fixé certaines contraintes pour développer un système permettant le maintien à domicile. Notamment, le système doit être basé sur une technologie réaliste, c’est-à-dire utilisable. Nous voulions, entre autres, développer des algorithmes qui puissent fonctionner sur des machines à bas coût. Pour cela, nous proposons des algorithmes simples et peu coûteux en temps de calcul et en mémoire, afin de pouvoir les insérer dans n’importe quelle machine. Pour être utilisable, l’équipement nécessaire ne doit pas être trop encombrant pour pouvoir être installé à domicile. La contrainte d’utilisation de capteurs ambiants, plutôt que de capteurs portés par la personne, a également été posée. L’idée est de laisser libre la personne sans qu’elle n’ait besoin de penser à porter le système ou à l’activer. Le choix final s’est donc porté sur le capteur Kinect de Microsoft qui est une caméra de profondeur à bas coût.

3.2 Contributions

Dans l’organisation de ce mémoire, nous distinguons deux contributions. La première porte sur l’analyse du mouvement de la personne et la seconde sur l’analyse de l’activité. Ces deux contributions pourront permettre la conception d’un système d’évaluation et de sécurisation de la personne à son domicile. L’analyse du mouvement et la détection des activités d’une personne sont deux thèmes de recherche traités dans la littérature séparément, car chacun a ses propres difficultés faisant appel à des compétences différentes. Nous montrons dans cette thèse que les deux problématiques peuvent être abordées par l’analyse d’un ensemble de paramètres simples extraits d’images de profondeur.

La partie I présente les systèmes existants pour d’un côté mesurer le degré de fragilité à l’hôpital et de l’autre détecter les chutes au domicile. En observant les systèmes existants et les contraintes que nous nous sommes fixées, nous expliquons les raisons pour lesquels nous avons choisi les caméras de profondeur pour développer le système.

Dans une deuxième partie, nous présentons les différents algorithmes conçus à partir des informations obtenues avec la caméra. Nous y apportons, au chapitre 3, un état de l’art non exhaustif des méthodes utilisées en vision pour détecter et suivre une personne au cours du temps. Puis, au chapitre 4, nous développons nos algorithmes de détection et de suivi de la personne, fondés sur des méthodes simples que sont l’extraction de fond et le suivi de la silhouette de la personne. Ces algorithmes sont communs pour l’analyse de la marche et la détection d’activité. Plus précisément, les deux contributions sont fondées sur l’analyse des mêmes paramètres, ex-

traits des images de profondeur fournies par la caméra, qui sont le centre de masse estimé de la personne et la distribution verticale des points de la silhouette. Nous développons dans deux chapitres distincts, les algorithmes propres à chacun des deux axes de recherche. Le chapitre 5 de ce mémoire présente les aspects concernant la marche des personnes. Ce chapitre introduit, dans un premier temps, un certain nombre de notions et de techniques, notamment dans le domaine médical, auxquelles nous faisons référence. Nous présentons les différents paramètres qu'il est possible d'analyser pour évaluer la marche d'une personne. Puis, nous présentons les paramètres que nous avons sélectionnés (longueur de pas, cadence et vitesse de marche) et les algorithmes d'extraction de ces paramètres. Cette contribution concerne la mesure de paramètres de la marche comme indicateur de l'évolution de la fragilité chez les personnes âgées. Le chapitre 6 de ce mémoire est consacré au développement d'un détecteur d'activité. Il repose sur l'utilisation du formalisme des modèles de Markov cachés, défini en annexe A. Dans ce chapitre, nous développons les différents modèles construits pour répondre au problème de la détection d'activité au domicile des personnes. La détection de l'activité apporte des informations à la fois pour la sécurisation et l'évaluation de la personne âgée. D'un côté, le travail présenté contribue à la problématique de sécurisation en détectant les chutes et les comportements anormaux. De l'autre, il contribue à la problématique d'évaluation de la perte d'autonomie permettant une analyse au long cours de l'évolution des habitudes de la personne.

Les contributions des chapitres 5 et 6 sont ensuite évaluées indépendamment à la partie III de ce mémoire. Dans une première expérimentation en laboratoire nous montrons, sur des personnes jeunes, que la précision des paramètres de la marche extraits avec notre algorithme est similaire aux valeurs extraites d'un tapis actimétrique, pris comme mesure de référence. D'autres expérimentations en laboratoire ont permis de constater que certains modèles de Markov cachés fournissent une classification robuste des activités, aussi bien lorsque la personne est entièrement visible que partiellement visible. Des premiers tests ont également été effectués avec des personnes âgées en situation réelle (à l'hôpital et au domicile). Les résultats ont montré que, malgré la variabilité des trajectoires du centre de masse des sujets, l'extraction des paramètres de la marche et la reconnaissance des activités sont possibles.

Enfin, la partie IV présente, dans un premier chapitre, l'association des deux contributions dans un système unique aidant à la sécurisation et à l'évaluation pour le maintien à domicile des personnes âgées. Plus précisément, nous intégrons l'algorithme de détection d'activité et celui de l'extraction des paramètres de la marche dans un seul système. Ainsi, selon l'activité détectée par l'algorithme, celui-ci enverra une alerte aux secours en cas de chute, comptabilisera le temps passé dans des activités passives (allongé ou assis) et actives (marcher ou être debout) ou analysera les paramètres de la marche quand la personne sera détectée comme marchant. Nous développons également une méthode pour cartographier les habitudes des personnes âgées, c'est-à-dire apprendre les habitudes les plus fréquentes dans les différentes zones du logement. Cela permettra de détecter les évolutions des habitudes de la personne et donc de détecter les anomalies. La carte est construite à partir de l'algorithme détectant les activités de la personne. La carte intègre aussi l'information concernant la vraisemblance permettant de distinguer les moments où la personne est entièrement ou partiellement visible. Dans le dernier chapitre, nous proposons d'autres perspectives pour le système. Nous pensons que ce système peut être utilisé avec d'autres objectifs que le maintien à domicile, comme différencier les personnes ou encore prévenir les troubles cognitifs et les risques d'hospitalisation.

L'apport de ce travail de thèse se situe à la fois sur le plan méthodologique et technique. Sur le plan méthodologique, la particularité de ce travail est qu'il est guidé par l'approche expérimentale. Les différentes décisions ont été prises avec l'objectif de développer un système correspondant au « réalisme du terrain », qui une fois construit, pourrait être expérimenté au domicile des personnes

âgées. Cette nécessité de développer des approches expérimentées et évaluées, au regard des contraintes particulières que pose l'environnement d'une personne âgée, a eu des répercussions sur le plan technique et sur les choix pris pour concevoir le système. L'objectif visé était de construire un système simple en vue des expérimentations futures. Nous montrons donc qu'en utilisant des indicateurs simples, que sont la trajectoire du centre de masse de la personne et la distribution verticale des points de sa silhouette, extraits à partir d'algorithmes utilisés en vision, il est possible de créer un système à haut pouvoir de décision. A partir de ces deux indicateurs, il est possible de sécuriser la personne dans son environnement, en détectant ses chutes et ses comportements dangereux, et d'évaluer son degré de fragilité en analysant sa marche et ses activités.

Première partie

Approches pour la mesure de la fragilité et la détection des chutes

Introduction

Nous avons posé les objectifs de ce travail, à savoir développer un système permettant de détecter les chutes et de mesurer le degré de fragilité d'une personne. Dans cette première partie, nous recensons les différentes approches existantes pour traiter ces deux questions. Les systèmes détectant la chute et analysant la fragilité des personnes sont souvent conçus pour résoudre l'un ou l'autre des deux problèmes. C'est pourquoi, nous avons construit l'état de l'art en deux chapitres.

Le premier chapitre présente les différentes approches pour évaluer le degré de fragilité d'une personne. Les dispositifs, décrits dans ce chapitre, sont des dispositifs utilisés dans le milieu médical. Les médecins analysent la fragilité, notamment, à travers les paramètres de la marche et leur évolution. Certaines modifications sont facilement perceptibles à l'œil. Les tests cliniques peuvent alors être suffisants. Les tests cliniques sont les premiers outils des médecins pour analyser l'état général de la marche d'une personne. D'autres modifications de la marche requièrent des instruments de mesure, comme c'est souvent le cas pour évaluer les progrès réalisés en rééducation. Nous retrouvons des systèmes avec capteurs ou sans capteurs portés par la personne. Dans la catégorie des capteurs portés, les dispositifs sont nombreux, allant du simple accéléromètre au capteur plus évolué comme l'électromyographie, en passant par des systèmes de chaussures équipées de capteurs de pression ou encore des systèmes de caméras avec capteurs portés par les personnes. Les dispositifs, dans la catégorie des systèmes sans capteurs portés, sont les tapis roulants, les tapis actimétriques, ou encore les plateformes de force.

Dans un deuxième chapitre, nous recensons les systèmes utilisés pour détecter la chute des personnes au domicile. Le découpage à l'intérieur de ce chapitre est effectué selon le type de capteurs utilisé : capteurs embarqués portés par la personne ou capteurs ambiants. Dans la catégorie des systèmes embarqués sur la personne, nous citons les capteurs tels que les accéléromètres ou les goniomètres. Concernant les systèmes ambiants, nous citons les capteurs environnementaux, comme les capteurs infrarouges de mouvement ou les détecteurs d'ouverture. Les systèmes à base de caméra appartiennent également à cette catégorie. Et enfin, plus récemment, des dispositifs placés au sol apparaissent.

Chaque système décrit a plus ou moins d'avantages et d'inconvénients par rapport à ce que nous souhaitons développer. Les systèmes adaptés uniquement à une utilisation en cabinet médical (les systèmes à base de capteurs embarqués), fournissant peu d'information (comme les accéléromètres) ou encore ayant un prix élevé, ne correspondent pas à nos objectifs.

Chapitre 1

Panorama des approches pour l'analyse de la fragilité à l'hôpital

Sommaire

1.1	L'œil expert à travers les tests cliniques	13
1.2	Outils d'analyse du mouvement	14
1.2.1	Systèmes avec capteurs embarqués sur la personne	14
1.2.2	Systèmes sans capteur embarqué sur la personne	16
1.3	Conclusion	16

La fragilité des personnes est évaluée essentiellement par le corps médical à travers une analyse de leurs capacités motrices. Les médecins analysent la marche des personnes à travers des tests cliniques. Ils peuvent, également, utiliser des outils de mesure pour quantifier la marche.

1.1 L'œil expert à travers les tests cliniques

L'analyse de la marche à travers l'œil humain est une méthode souvent utilisée dans le monde médical, car peu coûteuse. Les médecins observent une éventuelle pathologie ou fragilité chez les personnes âgées en s'aidant d'outils tels que la vidéo, le chronomètre, le mètre et les tests cliniques. De nombreux tests ont été mis en place pour aider l'évaluation faite par le médecin. Aucun consensus n'a été posé sur le choix des tests à effectuer. Le médecin peut donc choisir parmi un large échantillon de test. La liste suivante n'est pas exhaustive, mais elle permet de citer quelques tests couramment utilisés durant la consultation des patients en cabinet :

- le test des « dix mètres de marche » où le médecin calcule le nombre de pas et/ou chronomètre le temps pour faire cette tâche ;
- le test « up and go » où le médecin calcule le temps mis par la personne pour s'asseoir, se lever, marcher sur 3 m et revenir s'asseoir ;
- le test de « durée de l'appui unipodal » où le médecin calcule le temps durant lequel la personne tient l'équilibre sur un pied sans se tenir ;
- le test de la « vitesse de marche confortable » où le médecin calcule la vitesse du patient marchant « normalement » ;
- le test de « Tinetti » permettant, à l'aide d'un questionnaire et de petits exercices, de juger de la qualité de la marche et de l'équilibre.

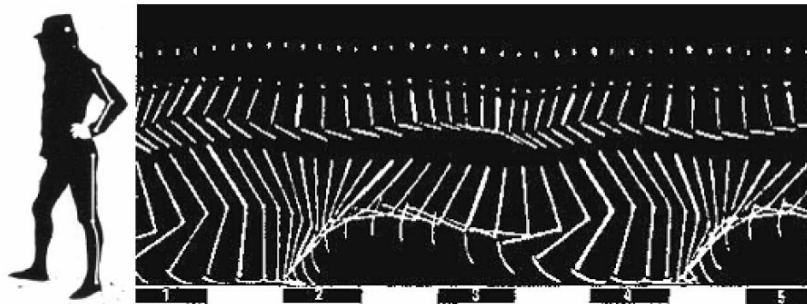


FIGURE 1.1 – Représentation de la marche humaine en diagramme bâton par Marey.

Ces tests permettent de fournir une évaluation de la qualité de la marche et de l'équilibre de la personne pour ainsi évaluer son état. Chaque médecin peut avoir un regard différent sur les résultats. Pour pallier au manque d'objectivité, d'autres techniques existent.

1.2 Outils d'analyse du mouvement

Les systèmes de mesure du mouvement sont utilisés dans divers domaines et pour divers objectifs. Ils peuvent être utilisés :

- pour étudier la marche (normale ou pathologique) de la personne dans le domaine médical ou biomécanique ;
- pour analyser le corps en activité dans le domaine du sport ;
- pour étudier l'influence de différents dispositifs sur la marche en ergonomie ;
- pour analyser la marche humanoïde dans le but de la comprendre et la reproduire dans le cadre de la robotique ;
- pour simuler la marche et l'intégrer dans des personnages virtuels dans le domaine de l'animation et la simulation de synthèse.

L'objectif commun des différents domaines est de quantifier la marche, de l'analyser dans un contexte spatio-temporel. Dès la renaissance, le sujet intéresse les scientifiques. Léonard de Vinci, Galilée et Newton analysaient déjà la marche. Mais le premier instrument de mesure naquit avec l'apparition de la chronophotographie. Cette technique consiste à prendre des photographies lors d'un mouvement pour le décomposer et le représenter sous la forme d'un diagramme de bâton comme illustré par la figure 1.1 [Marey, 1894]. En 1878 Eadweard Muybridge et en 1882 Étienne-Jules Marey utilisèrent les premiers cette technique. Muybridge analysa le mouvement du cheval et Marey le mouvement de l'oiseau en vol. Ces travaux ont constitué les prémisses du développement de nombreux outils de mesure du mouvement humain. Aujourd'hui, il existe de nombreux systèmes pour mesurer le mouvement. Nous les séparons en deux catégories : les systèmes nécessitant de placer des capteurs sur la personne et ceux sans capteurs embarqués.

1.2.1 Systèmes avec capteurs embarqués sur la personne

Différents capteurs, à poser sur le corps de la personne, existent permettant de fournir plus ou moins d'indications sur la marche. Le capteur de plus bas niveau, car apportant peu d'information, est le podomètre, comptant simplement le nombre de pas effectués par la personne lorsqu'elle marche. Cet outil peut être utile également, en dehors du cabinet du médecin, pour mener des recherches. Par exemple, en 2002 Silva *et al.* [Silva *et al.*, 2002] ont réalisé une étude



FIGURE 1.2 – Le système Walkmeter© est inspiré du locomètre de Bessou.

sur le nombre de pas par an effectués par des personnes et ont cherché des spécificités selon la catégorie de personne (homme, femme, personne obèse, etc.).

Les capteurs de mouvements (les capteurs infrarouges passifs) fournissent également un seul indicateur qui est la vitesse de marche. Un exemple d'utilisation peut être donné avec l'étude en 2010 réalisée par Hagler *et al.* [Hagler *et al.*, 2010]. Ils ont eu l'idée d'installer 4 capteurs infrarouges (PIR) au plafond du domicile des personnes âgées. Le but était d'observer un éventuel déclin à travers la vitesse de la marche et ainsi poursuivre l'analyse en dehors du cabinet du médecin .

Le locomètre de Bessou est l'instrument représenté à la figure 1.2. Il est composé d'une boîte de laquelle deux câbles sortent, pour être attachés aux deux talons. Ainsi, ce système fournit la mesure des déplacements longitudinaux des câbles. Plusieurs paramètres de la marche en sont déduits comme la longueur des pas, la fréquence du pas, la durée des appuis, etc. Ce système est limité par la longueur des câbles. Un exemple d'utilisation pour les personnes hémiplegiques est présenté dans l'article de Condouret *et al.* en 1987 [Condouret *et al.*, 1987].

Il existe également des chaussures munies de capteurs de pression permettant de récupérer les positions spatio-temporelles des pieds. A partir de ces positions, des valeurs telles que la vitesse de marche, les longueurs de pas peuvent être calculées.

Les accéléromètres, les gyroscopes ou les goniomètres ([Bourke *et al.*, 2007], [Wu, 2000]) sont des capteurs posés sur la jambe de la personne. Ils permettent d'analyser la marche et plus précisément l'accélération du corps ou l'angle des différents segments de la jambe lors de la marche. Ces systèmes sont posés sur les parties molles de la jambe donc il est nécessaire de vérifier qu'ils ne bougent pas au cours de la séance. De plus, le positionnement de ces capteurs varie selon la personne qui les pose.

Les systèmes optoélectroniques à marqueurs actifs s'insèrent dans cette catégorie (comme le système Selspot). La personne marche dans le champ de vision des caméras qui captent la lumière émise par les capteurs, posés sur la personne à différents endroits du corps. Cette technique permet d'obtenir très précisément les trajectoires spatiales et ainsi d'en déduire un certain nombre de paramètres comme les longueurs de pas, les angles des différents segments du corps, etc. Chaque capteur est alimenté par un câble, ce qui limite les déplacements de la personne. Il existe également des systèmes optoélectroniques à marqueurs passifs permettant aux capteurs de ne plus être reliés à des câbles d'alimentation. Les systèmes Vicon et Qualisys, par exemple, sont

des systèmes à marqueurs passifs. Ces systèmes consistent à positionner, en certains points clés du corps, des marqueurs réfléchissant un signal infrarouge émis par les caméras. Ainsi, la position 3D des différentes parties du corps est évaluée et permet, tout comme les systèmes à marqueurs actifs, de calculer les longueurs de pas, les durées, les angles entre les segments, etc. Les systèmes optoélectroniques restent des systèmes assez contraignants et coûteux. A chaque nouvelle installation du système, une phase de calibration est nécessaire pour définir un repère commun entre toutes les caméras et les renseigner sur leur position les unes par rapport aux autres. Comme tout système à capteurs embarqués, le positionnement des capteurs reste dépendant de leur pose et peut bouger au cours de la marche. De plus, certains capteurs peuvent durant la marche être occultés par les segments corporels. Pour analyser correctement le mouvement de la personne, six caméras au minimum sont généralement nécessaires. Ces dispositifs sont toutefois très utilisés dans différents secteurs, notamment en milieu hospitalier, pour la simulation de personnages virtuels, pour le sport, pour la biomécanique, etc.

L'électromyographie (EMG) est un système permettant d'enregistrer l'activité électrique musculaire en positionnant des électrodes sur le corps ou à l'intérieur du corps sur les muscles à étudier. Cette technique est assez controversée, car l'activité électrique enregistrée peut provenir de l'activité de muscles voisins et non pas de celui que l'on cherche à analyser.

1.2.2 Systèmes sans capteur embarqué sur la personne

Des systèmes sans capteurs embarqués sur la personne peuvent également être utilisés, la laissant libre lors de ses mouvements. Les plates-formes de force sont des outils sans capteurs embarqués sur la personne. Ces systèmes sont composés d'une simple plateforme (d'environ 60*60 cm) placée sur le parcours de marche. Ils permettent de mesurer les forces appliquées par le pied sur la plateforme lors de la marche.

Les tapis roulants instrumentés (figure 1.3) constituent également des systèmes permettant d'analyser la marche. Ils sont constitués de plateformes de force. L'avantage, comparé à la simple plateforme de force, est que le tapis analyse un grand nombre de pas successifs. La personne marche au rythme imposé par le tapis. La vitesse de marche de la personne n'est donc pas naturelle et ne peut pas être prise en compte pour juger de la qualité de la marche. D'autres paramètres comme la longueur de pas et les forces du pied sur le tapis peuvent être analysés.

Les tapis actimétriques (figure 1.4) permettent également d'enregistrer la pression plantaire exercée par le pied sur le tapis. La force appliquée par le pied est alors calculée, permettant de déduire des paramètres tels que la longueur de pas, la vitesse de marche, etc. Ce système laisse la personne libre de marcher mais il est limité par sa taille. Ce genre de tapis n'excède pas les 6 m de long.

1.3 Conclusion

Le milieu hospitalier est le premier endroit où l'on retrouve des systèmes analysant le mouvement des personnes dans le but d'évaluer le degré de fragilité d'une personne. La plupart de ces outils sont coûteux et nécessitent une expertise technique forte pour leur mise en œuvre. C'est pourquoi, chaque institut développe ses propres protocoles d'analyse, suivant les outils qu'ils ont à leur disposition.



FIGURE 1.3 – Le tapis roulant instrumenté de la société Biometrics.



FIGURE 1.4 – Le tapis actimétrique de la marque GaitRite.

Chapitre 2

Panorama des systèmes de détection de chutes au domicile

Sommaire

2.1	Systèmes avec capteurs embarqués sur la personne	19
2.1.1	Détecteurs de chutes automatiques	19
2.1.2	Systèmes d’alerte manuels	21
2.2	Systèmes ambiants	21
2.2.1	Capteurs environnementaux : la télé-assistance	21
2.2.2	Systèmes à base de caméra : la vidéo-vigilance	22
2.2.3	Capteurs au sol	22
2.3	Conclusion	22

Nous présentons dans ce chapitre les dispositifs existants pour détecter les chutes d’une personne. Le marché des systèmes de détection de chutes est un marché grandissant, principalement destiné à la surveillance au domicile des personnes. En recherche, le sujet de la détection de chutes est également très étudié. Nous avons regroupé les capteurs permettant de détecter les chutes en deux catégories, d’un côté les systèmes utilisant des capteurs à poser sur la personne et de l’autre les systèmes de capteurs ambiants.

2.1 Systèmes avec capteurs embarqués sur la personne

La première catégorie regroupe les systèmes à base de capteurs portés sur la personne. L’avantage de ces dispositifs est que bien souvent ils peuvent fonctionner à domicile et à l’extérieur.

2.1.1 Détecteurs de chutes automatiques

Les accéléromètres, les gyroscopes et les goniomètres sont des capteurs pouvant servir à détecter automatiquement les chutes des personnes âgées. Dans le domaine de la recherche, certains auteurs, comme Bourke *et al.* [Bourke *et al.*, 2007] et Wu [Wu, 2000], les ont d’ailleurs utilisés avec l’objectif de tester leur efficacité à détecter les chutes. Leur idée était de positionner ces capteurs sur la personne et d’observer la vitesse de mouvement du corps lorsque la personne réalisait différentes activités. L’hypothèse est que la vitesse est différente quand une personne chute, en comparaison à d’autres activités comme s’asseoir ou s’allonger dans un canapé.



FIGURE 2.1 – Système d’alerte de Vitalbase.

Ces capteurs ont été évalués en recherche et aujourd’hui, dans le commerce, certains dispositifs intègrent ce genre de capteurs. Des industriels, tels que Telecom Design avec VitalBase, intègrent un accéléromètre dans un dispositif à positionner au niveau de la ceinture ou au poignet comme le montre la figure 2.1. Ce système permet de détecter la chute à travers la perte de verticalité, l’accélération, le choc de la chute et enfin l’absence de mouvement. Les chutes brutales peuvent être détectées mais pas les chutes molles. Pour éviter les fausses alarmes, une phase de vérification est intégrée au système. Lorsqu’une chute est détectée, le système se met à vibrer durant 26 secondes et la personne doit, durant cette période, réaliser un mouvement ample du bras. Ce geste permet de bloquer l’envoi de l’alerte.

Le téléphone portable de l’entreprise suédoise Doro intègre également un accéléromètre. Ce téléphone détecte la chute à travers la perte de verticalité. Ce système de détection de chute doit être activé par la personne lorsqu’elle pense en avoir besoin, par exemple lorsqu’elle se retrouve seule dans son appartement.

La société Vigilio a mis en place un dispositif, détectant les chutes, à base d’un capteur accélérométrique à porter sur le torse de la personne (figure 2.2(a)). En cas de chute ou d’immobilisation prolongée (dans le cas des chutes molles par exemple), leur boîtier (figure 2.2(b)) directement connecté à une ligne téléphonique pourra alerter automatiquement un numéro qui a été pré-enregistré.



(a) Capteur accélérométrique à porter sur le torse.



(b) Boîtier d’alerte.

FIGURE 2.2 – Système de détection de chute de la société Vigilio.

D’autres capteurs peuvent également être utilisés notamment des capteurs sensoriels pouvant détecter la perte de conscience des personnes. La société Vivago a intégré ces capteurs dans une montre. L’idée est de détecter la chute des personnes due à un malaise. En utilisant ces capteurs couplés aux capteurs de température cutanée, l’entreprise a mis au point un système de détection de l’état d’inconscience des personnes et d’alerte des secours.

2.1.2 Systèmes d'alerte manuels

Dans le commerce, il existe également des systèmes ne détectant pas automatiquement la chute. Ces systèmes requièrent la participation de la personne. Les boutons d'alarmes, existant sous forme de collier ou bracelet à porter sur soi, jour et nuit, permettent à la personne d'appuyer dessus pour alerter d'une chute. La marque Intervox du groupe Legrand ou encore le groupe Biotel-Tunstall développent, entre autres, des médaillons à porter par la personne. L'idée est d'éviter les fausses alarmes et de permettre à la personne de se sentir rassurée d'avoir à tout moment un dispositif d'alerte.

Ces dispositifs doivent être portés pour fonctionner. C'est la cause principale de leur manque d'efficacité, car les personnes ne les portent pas, soit parce qu'ils sont perçus comme stigmatisants soit parce que la personne oublie tout simplement de les porter. Des études ont aussi montré que ces dispositifs pouvaient être inefficaces, même lorsqu'ils sont portés, car les personnes ne les utilisent pas forcément après une chute. Dans l'article de Wild *et al.* [Wild *et al.*, 1981], les auteurs ont mené une étude sur les personnes âgées de plus de 65 ans durant un an. Dans cette expérience, ils ont formé deux groupes, un groupe de « chuteurs » (constitué de personnes ayant déjà chuté une fois) et un groupe contrôle. Parmi les neuf chuteurs ayant un système d'alerte, seulement deux ont réussi à l'utiliser pour alerter après être tombés, l'un a tenté mais n'a pas réussi à l'utiliser et les six derniers n'ont même pas tenté de l'utiliser. L'incapacité d'appuyer sur le bouton peut être due à la perte de connaissance de la personne ou à son état de choc.

2.2 Systèmes ambiants

Cette catégorie de dispositif, basée sur des capteurs ambiants, ne fonctionne qu'à l'intérieur. Ces systèmes ont l'avantage de fonctionner automatiquement sans que la personne n'ait besoin de le déclencher.

2.2.1 Capteurs environnementaux : la télé-assistance

Les capteurs environnementaux permettent de rendre compte de l'activité d'une personne dans son lieu de vie. On peut compter le nombre de fois où la personne ouvre le réfrigérateur, le nombre de fois où elle s'est assise, etc. Ces différentes informations permettent d'en déduire l'activité d'une personne. Par exemple, les capteurs peuvent indiquer qu'une personne est inactive depuis longtemps ce qui révèle un comportement anormal.

Différents capteurs peuvent être utilisés. Les capteurs infrarouges de mouvement informent de la présence ou non d'une personne dans une pièce. Ainsi, il est possible de s'assurer que la personne est sortie de sa chambre au cours de la journée par exemple. Des détecteurs d'ouverture, posés dans des endroits tels que le réfrigérateur ou le robinet, permettent de s'assurer de l'utilisation des différents points d'intérêts.

La marque Intervox propose toute une gamme de capteurs parmi lesquels figurent les capteurs de mouvements et d'ouverture. Le but est de définir des tranches horaires où la personne doit être aperçue dans telle pièce ou doit ouvrir tel point d'intérêt. En cas de non conformité avec les événements prédéfinis, une alerte est déclenchée.

Senioralerte utilise également ce genre de capteurs (détecteurs de présence, d'ouverture, etc.). L'objectif de cette entreprise est d'apprendre les habitudes des personnes au cours du temps pour déceler des activités anormales.

Ces systèmes, en cas de détection d'un éventuel problème, déclenchent une alerte qui sera envoyée aux proches ou à une assistance s'assurant auprès de la personne qu'un réel problème

est apparu avant d'appeler les urgences.

2.2.2 Systèmes à base de caméra : la vidéo-vigilance

Les systèmes à base de caméra sont plus intrusifs dans la conscience des gens que les capteurs environnementaux mais ils permettent d'obtenir plus d'informations. Notamment, même si certains capteurs peuvent déduire, suite à un ensemble d'événements, qu'une chute s'est produite, ils ne peuvent pas réellement l'identifier comme telle. Les systèmes à base de caméra peuvent intégrer des algorithmes analysant l'activité d'une personne et surtout des algorithmes détectant la chute, avec beaucoup plus de discernement que ne peuvent le faire les détecteurs classiques.

Peu de systèmes de vidéo-vigilance sont commercialisés. En France, le système EDAO propose d'installer des caméras à domicile pour détecter les chutes des personnes en perte d'autonomie. Ce système semi-automatisé détecte des situations à risque (grâce à des algorithmes d'analyse) et prévient un opérateur qui, en visionnant les images, alerte ou non la famille ou les secours.

Le sujet de la détection de l'activité et des chutes reste un sujet de recherche. Le but est de développer des algorithmes d'analyse de l'activité toujours plus performants, ne commettant aucune fausse alarme et détectant chaque situation de chute ([Jansen *et al.*, 2007], [Rougier *et al.*, 2011], [Nait-Charif and McKenna, 2004], [Auvinet *et al.*, 2011b], [Anderson *et al.*, 2009]).

2.2.3 Capteurs au sol

L'intérêt de ces capteurs est qu'ils sont dissimulés, les personnes âgées peuvent donc les oublier plus facilement et ne pas se sentir surveillées en permanence. Le principe de ces capteurs posés au sol est qu'ils détectent les pressions plantaires exercées par les personnes. Ainsi, les algorithmes développés peuvent déduire certaines informations, savoir dans quelle pièce se trouve la personne au cours de la journée, s'assurer qu'elle ne soit pas allongée au sol, etc.

En recherche, la thématique des sols « intelligents » commence à se développer. Rimminen *et al.* ([Rimminen *et al.*, 2009], [Rimminen *et al.*, 2010]) décrivent dans leurs deux articles une technique pour suivre la personne (en utilisant entre autres un filtre de Kalman) puis pour détecter les chutes (en utilisant notamment des chaînes de Markov). Un appartement construit au sein d'INRIA à Nancy a été recouvert d'un réseau de capteurs au sol ([Pepin *et al.*, 2009]). Le sol est composé de dalles, mesurant 60×60 cm, qui ont été construites sur mesure. 100 dalles ont été installées dans cet appartement comme le montre la figure 2.3(a). Chaque dalle est équipée d'un accéléromètre, d'un magnétomètre et de quatre jauges de contraintes (des capteurs de force) positionnées à chaque angle comme précisé à la figure 2.3(b). Les dalles peuvent envoyer un message par le réseau ZigBee à d'autres dalles. L'un des objectifs, comme dans le travail de Rimminen, est de pouvoir détecter la chute des personnes parmi d'autres activités de la vie quotidienne. Rimminen travaille sur un sol instrumenté fondé sur la mesure du champ électrique.

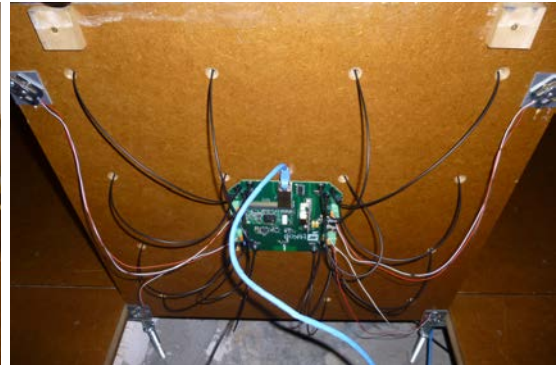
Un certain nombre de dispositifs commerciaux commencent à voir le jour. Pour n'en citer que quelques uns, nous pouvons mentionner les dispositifs proposés par Tarkett, Elsi et le SensFloor de la société Futur-Shape. Ces systèmes sont encore un peu coûteux et sont réservés à des installations dans les établissements d'hébergement pour personnes âgées dépendantes (EHPAD) ou dans des constructions neuves.

2.3 Conclusion

Les systèmes qui sont actuellement installés au domicile des personnes âgées détectent la chute ou les situations à risque. En dehors des boutons d'alerte, les systèmes de détection restent



(a) Appartement avec un réseau de dalles au sol.



(b) Dessous d'une dalle.

FIGURE 2.3 – Système d'analyse du mouvement par le sol d'INRIA à Nancy.

peu courants.

Conclusion

Dans cette partie, nous avons présenté un panorama des systèmes existant pour analyser la fragilité d'une personne et détecter sa chute. Ces deux fonctionnalités nous sont utiles pour développer un système de maintien des personnes âgées au domicile.

Les approches pour évaluer le degré de fragilité d'une personne sont surtout développées en hôpital ou en laboratoire. L'évaluation de la fragilité est réalisée à travers l'analyse de la marche. Les tests cliniques sont la méthode la plus simple à mettre en place et donc la plus répandue dans le milieu hospitalier. Les systèmes avec capteurs embarqués et sans capteurs embarqués peuvent être des techniques assez coûteuses et encombrantes mais plus précises que les tests cliniques pour quantifier le mouvement. Ces dispositifs ne correspondent pas à nos objectifs puisqu'ils ne sont pas adaptés pour être utilisés au domicile des personnes âgées dans le cadre d'un usage quotidien.

Les systèmes détectant la chute sont principalement conçus pour le domicile. Des capteurs à porter sur soi, comme des boutons d'alerte ou des bracelets détectant automatiquement la chute, existent mais sont peu utilisés, car ils sont stigmatisants pour les personnes. Des capteurs ambiants comme les capteurs environnementaux, les caméras ou les capteurs au sol sont conçus pour détecter automatiquement la chute au domicile.

Deuxième partie

Algorithmes pour l'évaluation et la sécurisation au domicile

Introduction

Nous développons dans cette partie les différents algorithmes et modèles que nous avons conçus pour suivre une personne dans le temps et l'espace, analyser son activité et sa marche. Nous exposons tout d'abord une vision globale du système proposé.

1 Vision générale

Nous exposons, ici, la vision que nous proposons d'un système pour favoriser le maintien à domicile des personnes âgées. L'architecture générale du dispositif est résumée par le schéma 3. L'objectif global est de sécuriser et d'évaluer la fragilité de la personne au quotidien à son domicile. Pour cela, nous prévoyons d'installer, dans chaque pièce de la maison, un capteur 3D couplé à un mini ordinateur de façon à traiter les données au domicile et ne pas transmettre d'images. Nous avons donc opté pour des algorithmes et modèles légers, ne nécessitant pas de gros calculs pour qu'ils puissent être implémentés dans des mini ordinateurs à bas coût et fonctionner en temps réel. A chaque information reçue du capteur, les algorithmes analysent la scène en détectant la personne et ses activités. Nous voulons identifier des activités de la vie quotidienne telles que s'asseoir, marcher, s'accroupir, se pencher, monter sur un obstacle (une chaise, une marche, etc.), s'allonger sur un lit ou sur un canapé, et des situations particulières comme chuter et être allongé au sol. Ainsi, à la fin de chaque journée, un bilan du temps d'activité et d'inactivité pourra être fait. Cet indicateur de l'évolution de la fragilité de la personne sera couplé à l'évaluation de l'évolution de sa marche. Les observations sur l'activité et la marche de la personne seront envoyées, selon le protocole de suivi, au médecin, à la famille et à la personne elle-même pour ainsi évaluer l'évolution jour après jour de son état. Le médecin pourra alors coupler ces informations avec les différents tests qu'il effectuera durant les consultations pour affiner son évaluation de la fragilité de la personne. D'une manière plus générale, le système pourrait apprendre les habitudes des personnes, comme l'heure à laquelle elle se réveille ou la vitesse à laquelle la personne marche. Ainsi, les changements d'habitudes pourront être utilisés comme des indicateurs de l'évolution de l'état général de la personne. La deuxième fonctionnalité du système serait de sécuriser la personne en alertant les secours en cas de chute. Le système pourrait également prévenir la famille en cas de détection de situations à risque telles que monter sur une chaise et s'accroupir, qui sont des activités pouvant conduire à une chute.

2 Choix des capteurs

Dans l'introduction de ce mémoire, nous avons indiqué quelques contraintes à prendre en compte dans le développement de notre système, notamment, il doit être réaliste par son coût et reposer sur des capteurs ambiants pour ne pas perturber la personne. Les systèmes avec capteurs embarqués sur la personne, tels que les accéléromètres, n'ont donc pas été retenus. Les capteurs

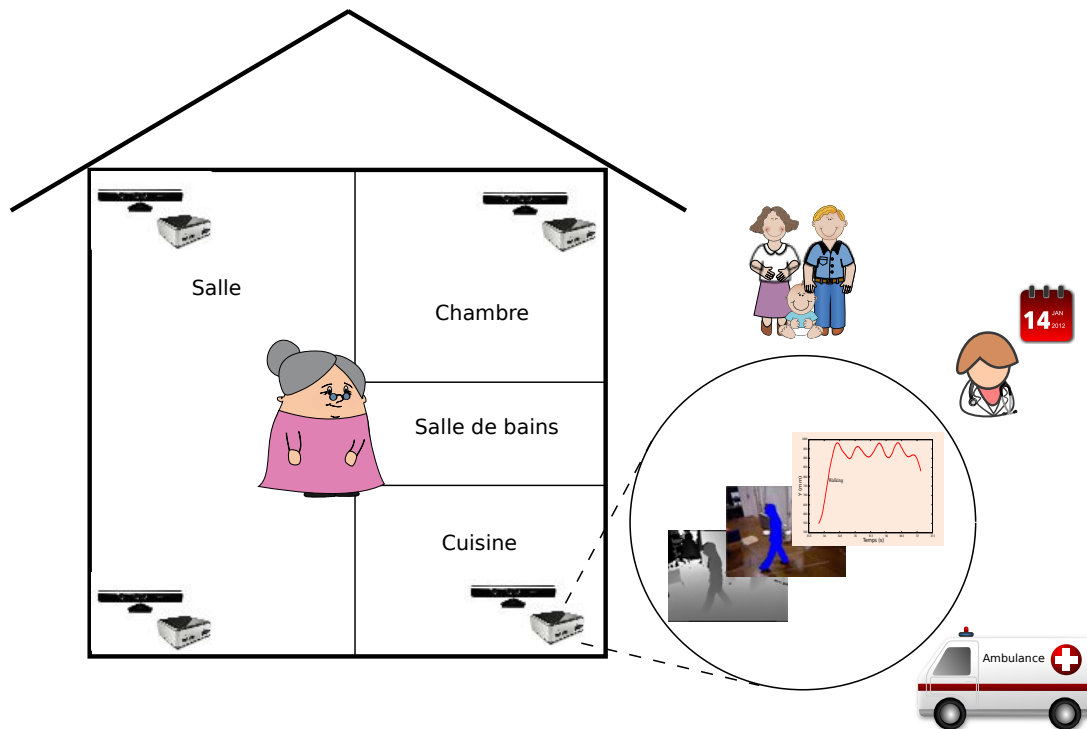


FIGURE 3 – Vision pour la conception d’un système permettant le maintien à domicile des personnes âgées.

au sol ne sont également pas retenus, car, aujourd’hui, ils n’apportent pas encore de solution facile à adopter dans un environnement déjà existant. Les capteurs environnementaux et les caméras correspondent au besoin. Il est possible d’installer un réseau de capteurs environnementaux au domicile de la personne sans la gêner et avec un coût peu élevé. Mais l’information sortant de ces capteurs est insuffisante pour ce que nous voulons faire. Ils permettent de détecter un grand nombre d’événements, comme l’heure à laquelle la personne a ouvert le réfrigérateur ou le nombre de fois où elle s’est levée de son fauteuil. Nous ne pouvons pas, en revanche, avoir d’information sur la qualité de sa marche. De plus, ces capteurs ne peuvent pas différencier les personnes. Ils ne peuvent pas savoir si la personne ouvrant le robinet est la personne à suivre, le conjoint, les visiteurs ou encore le personnel soignant.

Nous avons donc opté pour des caméras permettant, comme nous allons le montrer, aussi bien de détecter l’activité de la personne que d’analyser le degré de fragilité à travers l’évolution de sa marche. Les caméras RGB-Depth de Microsoft (les caméras Kinect) ont été choisies, car elles permettent de fournir une information de profondeur pour chaque objet en plus de l’image couleur. De plus, elles se trouvent facilement dans le commerce, sont à bas coût et fonctionnent la nuit.

3 Conception des algorithmes

La détection d’activité et l’analyse de la marche, à l’aide de caméras, nécessitent de concevoir au préalable un modèle d’extraction des personnes efficace et fiable. Les pixels extraits doivent ensuite être représentés sous une forme compacte permettant de suivre l’évolution dans l’image de la personne. Suivre une personne au cours du temps soulève plusieurs questions.

-
- Sous quelle forme représenter l’objet d’intérêt pour qu’il soit reconnu dans toutes les situations ?
 - A partir de quelles caractéristiques obtenir la représentation de l’objet ?
 - Comment extraire le ou les objets d’intérêts dans la scène ?
 - Quelle méthode est adaptée pour suivre l’objet au cours du temps ?

Dans le premier chapitre, nous présentons les méthodes utilisées pour traiter un flux vidéo. De nombreux articles traitent de la question du suivi d’objets (*Object tracking* en anglais) dans une vidéo. La liste des méthodes, que nous présentons dans ce chapitre, n’est donc pas exhaustive. Nous nous concentrons sur les méthodes permettant de détecter et suivre une personne au domicile. Suivre une personne dans une scène au cours du temps est un défi, car cela soulève de nombreux problèmes. Les difficultés rencontrées peuvent être dues à ses mouvements, à ses changements d’apparence (changement d’habit par exemple), à sa propriété non rigide (une personne est différente assise ou debout), aux changements de la scène (une chaise qui a bougé), aux occlusions pouvant cacher partiellement ou totalement la personne ou encore aux mouvements de la caméra. Dans ce premier chapitre, nous présentons la particularité de la caméra que nous utilisons. Les méthodes en vision ont été construites pour des caméras couleur classiques. L’apparition de la caméra fournissant une information de profondeur permet de faciliter un certain nombre de traitements qui auparavant étaient moins robustes.

Dans le deuxième chapitre, nous présentons les méthodes que nous avons implémentées pour détecter et suivre une personne. Nous présentons, tout d’abord, la méthode de soustraction du fond que nous avons identifiée comme peu coûteuse en temps de calcul et robuste si elle est faite à partir de l’image de profondeur renvoyée par la caméra. Puis, nous montrons qu’à travers cinq caractéristiques, extraites des images de profondeur, nous pouvons représenter une personne et la suivre au cours du temps. Ces cinq caractéristiques sont : le centre de masse, la vitesse verticale, la vitesse horizontale, la dispersion verticale et le point maximum. Ces caractéristiques ont été choisies, car elles sont faciles à extraire et peuvent être suivies en temps réel.

Dans le troisième chapitre, nous abordons l’algorithme permettant l’analyse de la marche d’une personne. L’idée est d’observer au cours du temps les paramètres de la marche d’une personne, comme indicateur de l’évolution de sa fragilité. Dans un premier temps, nous allons définir et décrire la marche. Nous verrons, entre autres, les différentes phases qui caractérisent la marche humaine et les indicateurs permettant de l’analyser. Nous verrons comment une analyse de différents paramètres spatio-temporels (longueurs de pas, vitesse de marche, etc.), cinématiques (angles formés par les différentes parties du corps) et dynamiques (les forces de réaction du pied, etc.) permet de révéler en grande partie les capacités physiques d’une personne. Pour notre étude nous nous sommes restreint à la seule mesure des paramètres spatio-temporels. Pour extraire ces paramètres, nous nous basons sur le centre de masse de la personne et sur la détection des maxima locaux sur l’axe vertical. La précision des paramètres de la marche est évaluée au chapitre 7 de la prochaine partie.

Le quatrième chapitre est consacré au développement d’un algorithme détectant l’activité d’une personne. Nous avons vu en introduction, l’importance de détecter au plus vite les chutes, pour éviter à la personne de rester au sol trop longtemps et pour diminuer le risque de séquelles physiques et psychologiques. La personne peut avoir peur de rechuter et, ainsi, minimise ses activités. En ce sens, la chute peut accélérer la perte d’autonomie. Cet axe de recherche a donc un double objectif. Il permet, d’un côté, de sécuriser la personne en détectant les chutes et les activités à risque, comme monter sur une chaise. Et, d’un autre côté, il permet d’évaluer le degré de fragilité de la personne et de prévenir la perte d’autonomie en détectant les comportements inhabituels, comme l’immobilité prolongée. Pour détecter l’activité d’une personne, nous avons choisi d’utiliser le formalisme des modèles de Markov cachés (MMC). Ce sont des modèles qui

permettent d'estimer un état caché à partir d'une séquence d'observations. Nous les définissons plus largement en annexe A. Parmi les nombreuses méthodes de la littérature, les MMC ont été souvent utilisés dans le cas de la reconnaissance d'activité. Les taux de reconnaissance d'activité sont très importants avec ce type de méthodes. Cependant, bien souvent, les algorithmes utilisent un grand nombre de caractéristiques pour déterminer l'état caché. Notre objectif est de construire des modèles de Markov cachés en prenant des caractéristiques simples à extraire et peu nombreuses, tout en conservant un bon niveau de classification des activités. Les caractéristiques, que nous prenons en compte, sont la position du centre de masse, la vitesse verticale et horizontale du centre de masse, la dispersion verticale de la silhouette et le point le plus haut du corps. Nous construisons plusieurs modèles pouvant tous correspondre à notre problème. Les différences entre les différents modèles portent sur le nombre d'activités que le modèle peut détecter (entre 8 et 9) et sur le nombre de caractéristiques pris en compte (entre 3 et 5). Nous avons également construit des modèles où nous fixons à la main certains paramètres et d'autres n'ayant pas de connaissance *a priori*. Nous avons voulu construire plusieurs modèles pour déterminer si, en prenant un nombre restreint de caractéristiques, le modèle peut discriminer les activités correctement ou s'il est nécessaire d'ajouter des caractéristiques pour augmenter le taux de bonnes classifications. L'évaluation expérimentale des résultats est présentée au chapitre 8 de la partie III.

Chapitre 3

Détecter et suivre une personne avec une caméra : état de l'art

Sommaire

3.1	Détection de la personne	33
3.1.1	Détecteur de points d'intérêt	34
3.1.2	Soustraction du fond	34
3.1.3	Segmentation	35
3.1.4	Apprentissage supervisé	35
3.2	Suivi de la personne	35
3.2.1	Modèles d'objets animés	35
3.2.2	Suivi de l'objet	37
3.3	Caméra de profondeur	40
3.4	Conclusion	41

Ce chapitre se base, entre autres, sur le livre de Russell et Norvig [Russell and Norvig, 2010] et sur une étude réalisée en 2006 par Alper Yilmaz *et al.* [Yilmaz *et al.*, 2006] recensant et classifiant les différentes méthodes utilisées dans la littérature, pour détecter et suivre une personne. Toutes les méthodes en vision ne sont pas abordées dans ce chapitre. Cet état de l'art est centré sur l'extraction d'une personne d'une scène avec une caméra fixe localisée en intérieur.

Les deux grandes étapes en « vision par ordinateur » sont développées dans deux sections séparées, dédiées à la détection d'objets animés et au suivi spatio-temporel de ces objets. Ensuite, nous abordons, dans une nouvelle section, la particularité du capteur de profondeur choisi qui, comme nous le verrons, simplifie fortement les traitements nécessaires à notre proposition.

3.1 Détection de la personne

La difficulté ici est de détecter dans la scène, vue par la caméra, la présence ou l'absence d'objet d'intérêt (la personne). Bien souvent la détection requiert plusieurs images. L'utilisation de seulement deux images consécutives peut provoquer de fausses détections. Par exemple, l'algorithme peut considérer qu'il y a un objet mobile alors qu'il s'agit juste d'une variation de la luminosité ou d'une ombre sur les deux images.

Dans l'état de l'art, plusieurs catégories de méthodes ont été proposées :

1. des méthodes détectant des points d'intérêts ;

2. des méthodes d'extraction de fond ;
3. des méthodes de segmentation de l'image ;
4. des méthodes d'apprentissage supervisé.

3.1.1 Détecteur de points d'intérêt

Le principe, ici, est de détecter des points d'intérêt dans l'image comme observé à la figure 3.1(b). L'avantage de ces méthodes est qu'elles sont robustes aux changements de lumière et indépendantes du champ de vision de la caméra. Le principe repose sur l'extraction des parties saillantes des objets à partir des variations d'intensité. Le détecteur de Moravec [Moravec, 1979], le détecteur de Harris [Harris and Stephens, 1988] et le détecteur SIFT [Lowe, 2004] (*Scale-invariant feature transform*) reposent sur ce principe. Ces détecteurs ont comme particularité d'être invariants aux mouvements de la caméra et des personnes.

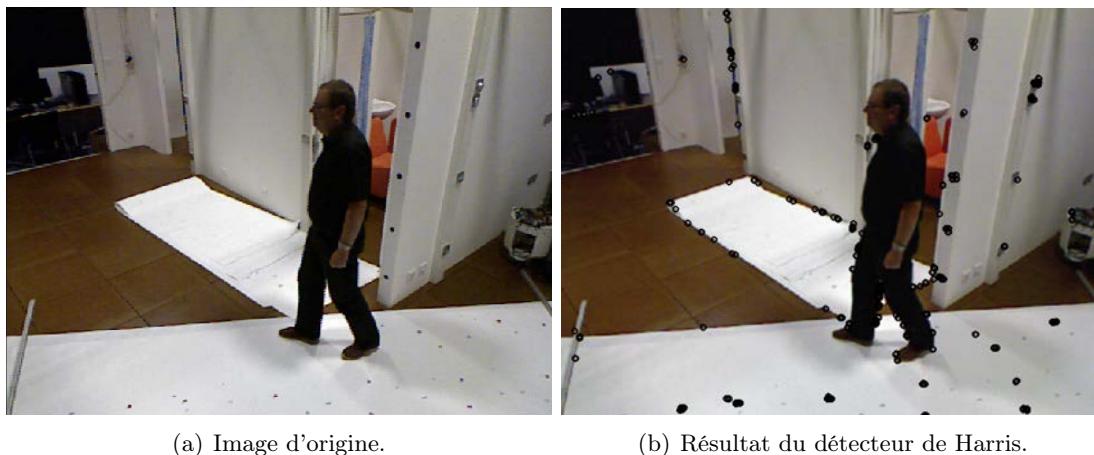


FIGURE 3.1 – Détection des points d'intérêts dans une scène.

3.1.2 Soustraction du fond

Les méthodes dans cette catégorie consistent à obtenir une représentation de la scène sur plusieurs images. Les objets mobiles recherchés sont détectés lorsque les images fournies par la caméra sont suffisamment différentes de la représentation du fond. Les méthodes fondées sur ce principe sont plus communément regroupées sous le terme de « soustraction de fond » et ont comme objectif de séparer les objets mobiles du fond.

Deux méthodes sont couramment utilisées dans la littérature pour obtenir une représentation du fond, la méthode des moyennes mobiles (*Running average* en anglais) et le mélange de gaussiennes (*Gaussian Mixture Model* en anglais). Dans la méthode des moyennes mobiles, chaque valeur de pixel, la valeur correspondant à la couleur par exemple, est moyennée sur plusieurs images pour obtenir une représentation du fond. Les valeurs des pixels des nouvelles images sont alors comparées aux valeurs des pixels du fond appris pour identifier les pixels mobiles. Des cas d'utilisation peuvent être trouvés dans la littérature, notamment dans les articles de Jansen *et al.* [Jansen *et al.*, 2007], de Gómez-Conde *et al.* [Gómez-Conde, 2011] et de Auvinet *et al.* [Auvinet *et al.*, 2011b]. Dans la méthode de mélange de gaussiennes, la valeur de chaque pixel est modélisée par un mélange de gaussienne. Pour chaque pixel, les paramètres du modèle,

à savoir les moyennes et les variances, sont appris à partir des observations obtenues sur plusieurs images consécutives. Cette technique a été utilisée pour la première fois par Stauffer *et al.* [Stauffer and Grimson, 2000].

Une comparaison des deux techniques a été faite par Gómez-Conde *et al.* [Gómez-Conde, 2011]. La méthode des moyennes mobiles est plus facile à implémenter et moins gourmande en calcul. Mais dans leur article, ils ont remarqué que cette méthode est moins robuste aux changements de couleur que le mélange de gaussienne.

3.1.3 Segmentation

La segmentation correspond au processus qui permet de découper l'image en régions de pixels similaires. Les pixels ont chacun des propriétés, par exemple une certaine couleur, luminosité et texture. Les pixels appartenant à un même objet ont des valeurs aux propriétés similaires. L'idée est de repérer les changements notables dans les valeurs prises par ses propriétés, pour ainsi découper l'image en région. Chaque région délimite les frontières entre les objets. Deux catégories de méthodes de segmentation existent, des méthodes détectant les frontières entre les objets [Martin *et al.*, 2004] et des méthodes détectant les régions [Shi and Malik, 2000]. Quelle que soit la méthode utilisée pour segmenter, la sélection des propriétés pour réaliser le regroupement est à définir en amont.

3.1.4 Apprentissage supervisé

L'idée est de regrouper en classe des échantillons aux caractéristiques similaires. Les nouveaux objets sont associés à la classe ayant les caractéristiques les plus proches. Le choix des caractéristiques est très important et doit permettre une bonne discrimination entre les classes. Cette méthode requiert un grand nombre d'exemples pour chaque classe. Pour classer les objets, différentes méthodes d'apprentissage existent comme les réseaux de neurones [Rowley *et al.*, 1998], *adaptive boosting* [Viola *et al.*, 2003] (AdaBoost), les arbres de décision [Grewe and Kak, 1995] et SVM [Papageorgiou *et al.*, 1998] (*Support Vector Machines*).

3.2 Suivi de la personne

Après avoir localisé la personne dans la scène (grâce aux méthodes de la section précédente), l'étape suivante est de suivre sa position dans le temps et dans l'espace. Pour suivre une personne, la première et la deuxième question de l'introduction de cette partie doivent être résolues, à savoir, sous quelle forme la personne sera représentée pour être suivie et à partir de quelles caractéristiques récupérer cette représentation.

Après avoir recensé les différentes façons de représenter l'objet d'intérêt, à savoir la personne, nous citerons quelques méthodes existantes pour le suivre selon la représentation choisie.

3.2.1 Modèles d'objets animés

Différentes représentations

Nous nous intéressons à la personne mais pour cette section, nous nous plaçons dans le cadre plus général de la représentation des objets. Les différentes manières de représenter l'objet d'intérêt peuvent être regroupées en deux catégories : les représentations fondées d'une part sur la forme de l'objet et, d'autre part, sur son apparence.

Concernant les représentations par rapport à la forme de la personne, certains auteurs la suivent à travers la silhouette (figure 3.2(a)) et les contours (figure 3.2(b)). La silhouette représentant la région à l'intérieur des contours. Le corps peut être représenté par des points (figure 3.2(c)) ou même par un seul point (figure 3.2(d)), par exemple, par le centre géométrique de la personne [Rougier *et al.*, 2011]. Pour représenter la personne, certains l'englobent dans une forme géométrique primitive, comme une ellipse (figure 3.2(e)) ou un rectangle (figure 3.2(f)), couramment appelée un « blob ». Anderson [Anderson *et al.*, 2006], Jansen *et al.* [Jansen *et al.*, 2007] et Auvinet *et al.* [Auvinet *et al.*, 2011b] utilisent des blobs pour représenter la totalité du corps humain. D'autres techniques existent permettant de représenter plus finement le corps en distinguant les différentes parties telles que le torse, les jambes, les bras, la tête ou les pieds (figure 3.2(g)). Chaque partie du corps peut être modélisée à travers un cylindre ou une ellipse, comme Rougier *et al.* qui entourent la tête [Rougier and Meunier, 2010] et Deeb *et al.* qui s'intéressent à la tête et aux pieds [Deeb *et al.*, 2012]. Une dernière méthode de représentation, fondée sur la forme du corps, est fondée sur l'extraction du squelette (figure 3.2(h)). Par exemple, Zhen-Peng Bian *et al.* [Bian *et al.*, 2012] et Saboune [Saboune and Charpillat, 2005] extraient le squelette dans le but d'étudier les chutes des personnes âgées.

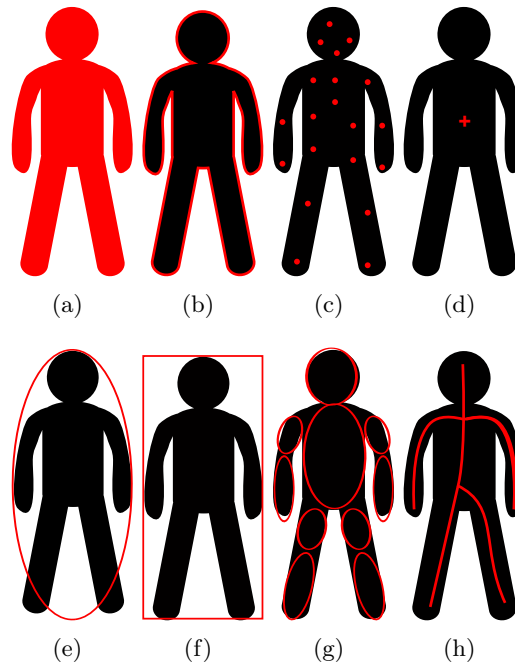


FIGURE 3.2 – Différentes représentations de l'objet à suivre : (a) silhouette, (b) contour, (c) points, (d) centre géométrique, (e) ellipse, (f) rectangle, (g) plusieurs ellipses et (h) squelette.

L'apparence est ce qui caractérise la ressemblance d'un objet. Deux objets se ressemblent s'ils ont la même apparence. L'apparence de certains objets varie peu c'est le cas, par exemple, d'une balle de tennis. Tous les objets n'ont pas une apparence aussi stable. Par exemple, les tables peuvent varier selon leur forme, leur couleur, leur taille et leur angle de vue. Les méthodes utilisées doivent donc être adaptées à l'apparence de l'objet à suivre. Dans notre cas, l'apparence d'une personne peut varier selon la façon dont elle est habillée, sa position, sa posture ou encore son angle de vue.

Ces différentes représentations peuvent être combinées. Bien souvent le choix de la représentation s'impose de lui-même selon l'objectif poursuivi.

Caractéristiques pour obtenir la représentation

L'algorithme doit se fonder sur certaines caractéristiques faciles à identifier de l'objet pour obtenir une représentation efficace (qui peut être sous la forme d'une ellipse, d'un point, d'un squelette, etc.). Plusieurs caractéristiques de l'objet peuvent être prises en compte.

Sa couleur Chaque pixel a trois valeurs, une pour le rouge, une pour le vert et une dernière pour le bleu.

Ses bords ou ses contours Les contours sont les lignes dans l'image où la luminosité est modifiée. En cherchant les différences de luminosité, il est possible de détecter les contours et donc de détecter des éléments importants de l'image, comme illustré à la figure 3.3(b).



(a) Image d'origine.

(b) Résultat de l'extraction des contours.

FIGURE 3.3 – Détection des contours dans une scène.

Son flot optique Le flot optique d'un objet révèle le mouvement d'un objet d'une image sur l'autre. Des groupes de pixels se déplaçant à la même vitesse et dans la même direction pourront alors être regroupés et associés à un objet particulier.

Sa texture La texture correspond au motif qui se répète sur la surface d'un objet et qui est perceptible visuellement. La texture doit être observée, non pas pour chaque pixel, mais par zone.

Les caractéristiques retenues pour représenter un objet sont choisies généralement en fonction du domaine, de l'objectif et du contexte. Des méthodes existent pour sélectionner automatiquement les caractéristiques discriminantes de l'objet. Les méthodes d'analyse en composantes principales ou la méthode Adaboost [Tieu and Viola, 2004] sont de bons exemples de méthodes ayant fait leurs preuves.

3.2.2 Suivi de l'objet

Le suivi d'un objet à travers une série d'images n'est pas un problème facile à résoudre, car plusieurs facteurs peuvent venir perturber le suivi comme les occlusions, les mauvaises détections de l'objet, les entrées et sorties du champ de vision de la caméra. La méthode, pour suivre un objet, dépend du choix pris pour représenter l'objet. La méthode ne sera pas la même si l'objet

est représenté par un ou des points, par une forme géométrique ou par sa silhouette, et en fonction des hypothèses faites sur sa dynamique.

Objet représenté par des points

Si l'objet est représenté à travers un ou plusieurs points, le principe de la méthode de suivi consiste à apparier les points d'une image aux points des images consécutives. Le problème consistant à faire correspondre les points, à travers plusieurs images, peut être résolu en utilisant une combinaison d'hypothèses :

- la position de l'objet dans des images consécutives ne doit pas énormément varier (3.4(a)) ;
- la vitesse de l'objet d'une image à l'autre n'excède pas un certain seuil, ce qui permet de supposer que le point à l'image $t + 1$ se trouvera dans un certain rayon comparé à l'image t (3.4(b)) ;
- la direction et la vitesse de l'objet ne changent pas énormément sur une succession de plusieurs images (3.4(d)) ;
- la vitesse des objets dans un petit voisinage est similaire (3.4(c)) ;
- l'objet en 3 dimensions est rigide, ainsi la distance entre deux points d'un même objet restera la même à la prochaine image (3.4(e)).

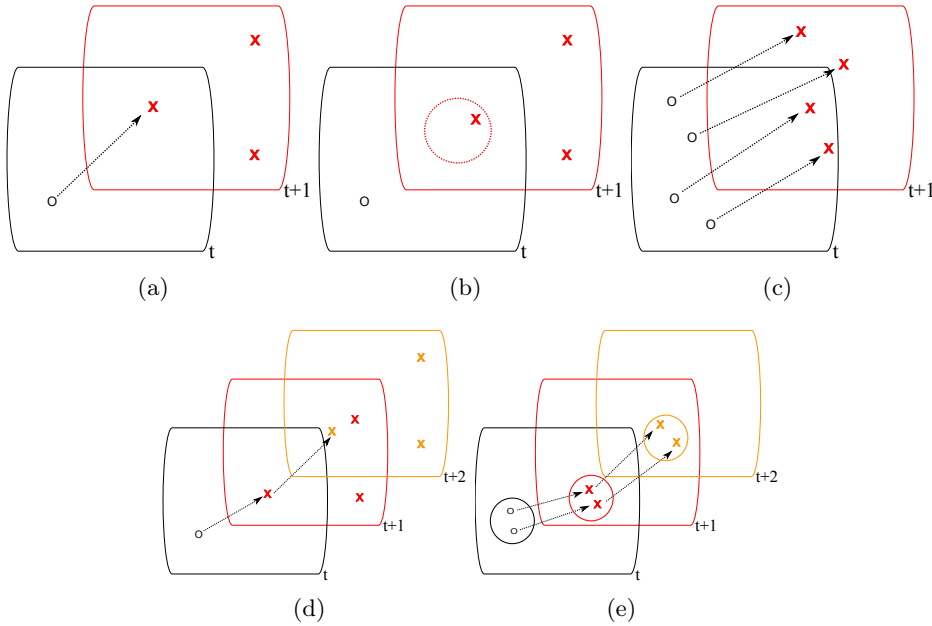


FIGURE 3.4 – Différentes hypothèses utilisées pour le suivi d'un objet : (a) proximité, (b) vitesse maximale, (c) mouvement similaire, (d) constance de la vitesse et de la direction, (e) rigidité de l'objet.

A partir de ces hypothèses, deux catégories de méthodes existent pour déterminer la correspondance entre les points de différentes images consécutives. Les approches déterministes consistent à évaluer et à minimiser un coût de correspondance entre les points de deux images consécutives. A chaque image, une solution unique est sélectionnée. Plusieurs auteurs traitent de ces approches. Nous pouvons citer en particulier Sethi et Jain [Sethi and Jain, 1987], Rangarajan et Shah [Rangarajan and Shah, 1991] et Veenman *et al.* [Veenman *et al.*, 2001]. Les approches statistiques (ou bayésiennes) consistent à estimer, à l'aide d'une mesure de l'incertain, la posi-

tion future de l'objet. L'idée est de supposer que l'observation peut être bruitée, ce qui induit de l'incertitude. A chaque image, ces approches fournissent un ensemble de solutions plus ou moins probables. Le filtre de Kalman (décrit à la section 4.2.2) et les filtres particuliers sont des approches bayésiennes souvent utilisées pour résoudre le problème de suivi d'une personne.

Objet représenté par des formes géométriques

L'objectif est de suivre l'objet en calculant le mouvement d'une image à l'autre, des formes géométriques qui l'entourent. Les mouvements des formes géométriques sont généralement sous la forme de transformations, comme la translation, la rotation et l'application affine. Un exemple de transformation d'une ellipse est présenté à la figure 3.5.

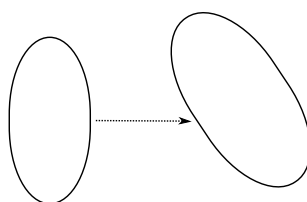


FIGURE 3.5 – Transformation paramétrique d'une ellipse.

Deux catégories de méthodes peuvent résoudre le problème du suivi du mouvement des formes géométriques d'un objet. La première catégorie se fonde sur l'hypothèse que la forme d'un objet change peu au cours du temps. L'algorithme définit un modèle d'apparence de l'objet, lorsqu'il apparaît dans la scène et recherche, au cours du temps, l'emplacement de ce modèle dans les images, pour ainsi suivre l'objet. Un exemple de modèle d'apparence est montré à la figure 3.6 représentant les différentes parties du corps par des rectangles. Chaque rectangle a ses propres dimensions et une certaine position par rapport aux autres. Plusieurs algorithmes ont été développés pour coupler le modèle à un élément de l'image, comme celui de Comaniciu [Comaniciu, 2002] et Jepson *et al.* [Jepson *et al.*, 2003]. L'autre catégorie de méthodes émet l'hypothèse que l'objet peut apparaître complètement différemment d'une vue à l'autre. Le modèle défini risque, dans ce cadre, de ne pas pouvoir correspondre à toutes les vues de l'objet et de ne pas le reconnaître dans une image, perdant alors son suivi. Le but est d'apprendre plusieurs vues de l'image avant de le suivre en direct, par exemple avec des SVM [Avidan, 2004].

Objet représenté par sa silhouette

La représentation par silhouette permet de suivre plus finement certaines formes complexes, comme la main, la tête, les épaules, que la représentation par forme géométrique. L'idée est d'extraire pour chaque image la silhouette de l'objet. La silhouette sera alors couplée à un modèle représentant la forme de l'objet ou représentant le contour. La forme de l'objet peut être suivie de la même manière que le suivi de modèle d'apparence du paragraphe précédent. La silhouette et son modèle associé sont recherchés à chaque image. La méthode de suivi de contour consiste à déplacer itérativement le contour initial dans l'image précédente pour le faire correspondre à celui de la nouvelle image. Le suivi de contour nécessite d'avoir un recouvrement des régions de l'objet dans les deux images consécutives.

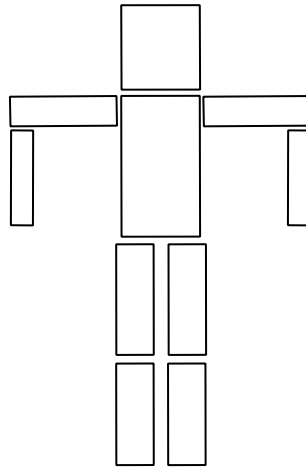


FIGURE 3.6 – Modèle d'apparence.

3.3 Caméra de profondeur

La caméra que nous utilisons est la caméra Kinect de Microsoft, comme montré à la figure 3.7. Elle a la particularité d'être un capteur actif fonctionnant dans l'infrarouge qui reconstitue, même de nuit, la profondeur de chaque pixel. Ses caractéristiques sont les suivantes. Elle fournit 30 images par seconde. Son champ de vision est de 57° à l'horizontale et 43° à la verticale. La profondeur peut être reconstruite jusqu'à 8 m avec une précision de quelques centimètres qui se dégrade au-delà de 4 m.

L'apparition de ce type de caméras simplifie les méthodes utilisées en vision. Les méthodes précédentes ont été développées pour des caméras fournissant une information d'intensité ou de couleur pour chaque pixel. La caméra de Microsoft, en plus de fournir l'information couleur, indique à quelle distance chaque pixel se trouve dans la scène par rapport à la caméra. Les caméras classiques ont l'inconvénient d'être dépendantes des changements de luminosité. Ainsi, la plupart des méthodes construites pour détecter les personnes sont conçues pour résoudre ce problème. La méthode de soustraction du fond, peu robuste avec des caméras classiques puisque le fond varie avec l'éclairage ambiant, est très efficace avec une image de profondeur [Shotton *et al.*, 2013].

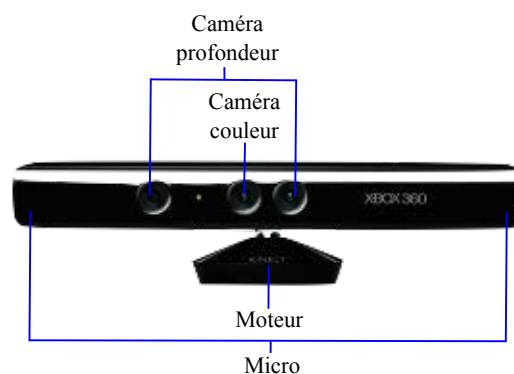


FIGURE 3.7 – Le capteur Kinect de Microsoft.

Microsoft a développé ses propres algorithmes de détection et suivi de la personne. Leur algorithme a été pensé pour être appliqué aux jeux où la personne est face à la caméra puisqu'à

l'origine cette caméra est un périphérique destiné à la console de jeux vidéo Xbox360. Ces algorithmes sont très robustes. L'algorithme est fondé sur le suivi d'un modèle de squelette. Le squelette continue d'être détecté, même si la personne est à moitié visible. Il repose sur une base d'apprentissage de 500 000 images. L'inconvénient de leur algorithme est qu'il fonctionne surtout si la personne est face à la caméra. De plus, si la personne est dans certaines positions, comme allongée au sol, leur algorithme ne détecte plus la personne. Il est peu robuste aux situations nouvelles n'appartenant pas à la base d'apprentissage. L'algorithme est propriétaire mais peut être utilisé via le SDK fourni par Microsoft. Nous avons fait le choix de ne pas utiliser le SDK de Microsoft pour développer une approche plus adaptée aux situations que nous souhaitons identifier tout en montrant que cet objectif peut être atteint par l'analyse de paramètres simples, sans passer par une reconstruction de la configuration 3D du squelette.

3.4 Conclusion

L'un de nos objectifs, comme nous l'avons mentionné, est de développer un système peu coûteux. En d'autres termes, le système doit pouvoir être installé sur une machine à bas coût. Les algorithmes doivent être simples et réaliser le traitement en temps réel. Les méthodes comme l'extraction du squelette, par exemple, sont des méthodes très précises mais coûteuses en temps. Une des méthodes les plus classiques pour détecter la personne en temps réel, et que nous avons sélectionnée pour développer notre système, est l'apprentissage du fond. Cette méthode est considérée comme l'une des moins robustes. Mais aujourd'hui, avec l'apparition des caméras nous fournissant l'information de profondeur, la représentation du fond peut être construite sur la profondeur de chaque pixel et non plus uniquement sur la couleur, la rendant ainsi plus robuste, car n'étant plus dépendante des changements de lumière. Nous avons choisi, pour suivre l'objet, une représentation en points du corps de la personne qui a l'avantage d'être une méthode légère.

Chapitre 4

Application à la détection et au suivi de personnes avec une caméra RGB-D

Sommaire

4.1	Détecter la personne	43
4.1.1	Extraction et apprentissage du fond	44
4.1.2	Extraction du sol et changement de repère	44
4.1.3	Extraction des points mobiles	46
4.1.4	Extraction des objets mobiles	47
4.1.5	Détection de la personne	49
4.2	Suivre la personne	49
4.2.1	Représentation de la personne	49
4.2.2	Suivi du centre de masse	51
4.2.3	Suivi de la distribution verticale	53
4.3	Conclusion	55

Dans ce chapitre, nous présentons les différentes étapes de traitement, allant de la détection de la personne, avec l'extraction des pixels mobiles, jusqu'au suivi d'une personne se déplaçant dans le champ de vision d'une caméra. Chaque étape de la détection et du suivi est schématisée sur la figure 4.1. Chaque encadré fait l'objet d'une sous section de ce chapitre. L'objectif final, de ce chapitre, est de suivre la personne à travers des paramètres simples à extraire. Nous avons choisi de suivre la personne à partir de son centre de masse et de la distribution verticale de sa silhouette. A partir de ces deux paramètres, nous pouvons déduire d'autres caractéristiques comme la vitesse verticale et la vitesse horizontale du centre de masse et le point maximum qui correspond au sommet du crâne lorsque la personne est debout.

4.1 Détecter la personne

Il s'agit dans cette section de détecter la personne à partir de l'image de profondeur fournie par la caméra. Dans une image de profondeur, chaque pixel représente un point dans l'espace 3D donné dans le système de coordonnées de la caméra. Les points du plan image peuvent être projetés dans ce système de coordonnées grâce aux paramètres optiques de la caméra. Ces paramètres sont des données constructeur propres à chaque caméra.

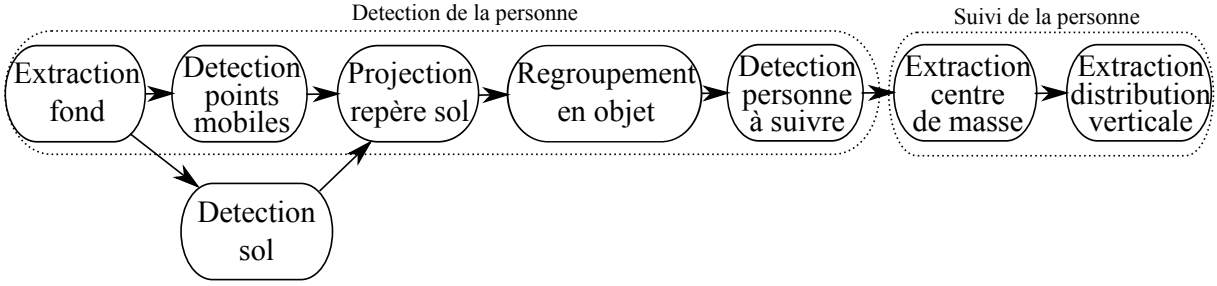


FIGURE 4.1 – Étapes pour détecter la personne du fond et la suivre dans le temps.

4.1.1 Extraction et apprentissage du fond

La première étape pour détecter la personne dans le champ de vision de la caméra est d'identifier les éléments fixes de la scène, que nous appelons le fond, pour pouvoir en déduire, par soustraction, les points appartenant à la silhouette de la personne. Tout d'abord, chaque pixel $F_0(x, y)$ représentant la carte de profondeur du fond au temps $t = 0$ est initialisé avec une valeur arbitraire nommée *MaxValue* et fixée à 12 mètres. Ensuite la méthode *running average* [Gómez-Conde, 2011] est utilisée pour apprendre le fond en moyennant, à chaque pas de temps, les distances pour chaque point de l'image. L'utilisation d'un poids α permet de donner plus d'importance au passé. La formule de mise à jour est la suivante :

$$F_t(x, y) = (1 - \alpha)F_{t-1}(x, y) + \alpha I_t(x, y)$$

avec F la matrice des distances cumulées, $I_t(x, y)$ l'image courante et α le poids.

Le poids α est fixé à 0,0001. Pour les points non reconstruits (comme dans le cas des surfaces réfléchissantes), la carte de profondeurs F n'est simplement pas mise à jour et reste égale à *MaxValue* de façon à permettre la détection des personnes devant les surfaces réfléchissantes.

Cette méthode permet de traiter le problème du fond non stationnaire. Quand un objet (comme une chaise) est déplacé, le fond est progressivement mis à jour.

Un certain temps est nécessaire pour apprendre le fond. Afin d'accélérer cet apprentissage, il est possible d'initialiser α à 1, pour le faire décroître ensuite progressivement jusqu'à 0,0001. Hormis cette phase initiale d'apprentissage du fond, mise en place pour faciliter les expérimentations, le système s'adapte automatiquement aux changements dans la scène.

4.1.2 Extraction du sol et changement de repère

L'objectif de cette étape est de permettre de faire, par la suite, des calculs indépendants de la position de la caméra. Les coordonnées 3D des points de l'image sont au départ calculées dans le système de coordonnées de la caméra dont l'origine est le centre optique. Afin d'être capable de mesurer le déplacement au sol de la personne et la hauteur des points extraits, nous souhaitons travailler dans un système de coordonnées aligné sur le sol. Pour effectuer ce changement de repère, nous avons besoin d'extraire le sol. Le système de coordonnées du sol est calculé juste après l'initialisation du fond, sans nécessiter l'utilisation d'une procédure de calibration spécifique.

Étape 1 : Redressement de la scène

Dans le repère de la caméra, nous désignons par X l'axe horizontal, Y l'axe vertical et Z la profondeur. Les points 3D du fond subissent tout d'abord une rotation autour de l'axe X en

utilisant l'angle γ d'inclinaison de la caméra obtenu à partir de son accéléromètre. Cette première transformation permet d'avoir un axe y proche de la verticale. L'équation de redressement des points P_i est la suivante :

$$P'_i = R_\gamma P_i$$

avec

$$R_\gamma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\gamma) & -\sin(\gamma) \\ 0 & \sin(\gamma) & \sin(\gamma) \end{pmatrix}.$$

Etape 2 : Sélection des points bas

Un nombre arbitraire de points sont ensuite sélectionnés parmi les points les plus bas. L'hypothèse est que la majorité de ces points appartiennent au sol. Nous prenons ainsi les 15 000 points les plus bas qui serviront à calculer l'équation du plan du sol.

Etape 3 : Calcul de l'équation du sol

Le sol est assimilé à un plan d'équation $ax + by + cz + d = 0$ dont nous calculons les paramètres a, b, c et d avec la méthode des moindres carrés. En rappelant que l'axe y est proche de la verticale suite à l'étape 1, nous pouvons considérer que b est non nul puisque l'axe y n'appartient pas au plan. L'ensemble des coefficients de l'équation du plan étant définis à une constante multiplicative près nous pouvons fixer $b = 1$. Chaque point P'_i sélectionné à l'étape 2 nous donne donc une équation du type :

$$x_i a + z_i c + 1d = -y_i.$$

Le système d'équations surdimensionné

$$Ax = B$$

où

$$A = \begin{pmatrix} x_1 & z_1 & 1 \\ \dots & \dots & \dots \\ x_N & z_N & 1 \end{pmatrix},$$

$$x = \begin{pmatrix} a \\ c \\ d \end{pmatrix}$$

et

$$B = \begin{pmatrix} -y_1 \\ \dots \\ -y_N \end{pmatrix}$$

peut être résolu avec la méthode des moindres carrés qui nous donne la solution

$$x = (A^T A)^{-1} A^T B$$

minimisant le résidu $r = B - Ax$.

Etape 4 : Construction du repère sol

Il s'agit enfin de construire le système de coordonnées du sol $(O_g, \vec{u}_x, \vec{u}_y, \vec{u}_z)$. Une origine arbitraire $O_g(0, y, 0)$ est définie telle que O_g appartienne au plan calculé précédemment, ce qui nous donne :

$$O_g = \begin{pmatrix} 0 \\ -d \\ 0 \end{pmatrix}.$$

Un premier axe est construit également arbitrairement en prenant un second point sur le sol de coordonnées $(1, y_x, 0)$ ce qui nous donne le vecteur directeur

$$\vec{u}_x = \alpha_x \begin{pmatrix} 1 \\ -a \\ 0 \end{pmatrix}.$$

L'axe Y_g est défini par le vecteur normal au plan

$$\vec{u}_y = \alpha_y \begin{pmatrix} a \\ 1 \\ c \end{pmatrix}.$$

L'axe Z_g est enfin construit en faisant le produit vectoriel entre u_x et u_y : $\vec{u}_z = \alpha_z(\vec{u}_x \wedge \vec{u}_y)$, soit

$$\vec{u}_z = \alpha_z \begin{pmatrix} -ac \\ -c \\ a^2 + 1 \end{pmatrix}.$$

Les nombres α_x , α_y et α_z sont des constantes de normalisation.

Projection d'un point dans le repère sol

Les points 3D $P(x, y, z)$ peuvent, après rotation suivant l'angle γ comme montré à l'étape 1, être projetés dans le système de coordonnées du sol $(O_g, \vec{u}_x, \vec{u}_y, \vec{u}_z)$ en réalisant les produits scalaires entre le vecteur $(O_g P')$ et les vecteurs directeurs du repère. Les coordonnées P'' de P dans le repère sol sont données par :

$$P'' = \begin{pmatrix} 1\alpha_x & -a\alpha_x & 0 \\ a\alpha_y & 1\alpha_y & c\alpha_y \\ ac\alpha_z & c\alpha_z & (a^2 + 1)\alpha_z \end{pmatrix} (P' - O_g)$$

où $P' = R_\gamma P$. Un exemple d'extraction du sol est montré sur la figure 4.2.

4.1.3 Extraction des points mobiles

Dans cette étape, nous cherchons à identifier les points mobiles de l'image, c'est-à-dire les points occupés par la personne. Pour cela, à chaque instant, le fond (l'image des distances du fond) est soustrait à l'image courante des distances, pour ne garder que les points ayant une distance plus courte dans l'image courante que dans l'image de fond.

Pour éliminer le bruit, c'est-à-dire les points détectés comme mobiles mais ne l'étant pas, un filtre Erosion-Dilatation est utilisé. Tout d'abord l'étape d'érosion permet d'éliminer tous les points mobiles proches d'un point considéré comme appartenant au fond. Ainsi, tous les



FIGURE 4.2 – L'extraction du sol est représentée en rouge.

contours des objets seront éliminés et le bruit également. La figure 4.3a) est un exemple de représentation des points mobiles et fixes où sera appliqué le filtre Erosion-Dilatation. Sur cette figure les cases en bleu représentent les points détectés comme mobiles et les points en vert représentent le fond. Le filtre érosion est appliqué à la figure 4.3a). Le résultat est montré en figure 4.3b). Toutes les cases bleues, ayant une case verte dans leur voisinage, deviennent vertes. Dans l'exemple, un voisinage de taille 1 est sélectionné, c'est-à-dire seuls les points juxtaposés sont pris en compte. La deuxième étape est le filtre dilatation consistant à transformer tous les points du fond, proches des points mobiles, en des points mobiles. A la figure 4.3b) est appliqué le filtre dilatation permettant d'obtenir la figure 4.3c). Chaque case verte juxtaposée à une case bleue devient bleue. Cela permet aux objets mobiles de récupérer leurs contours.

Dans notre algorithme la taille du voisinage est de 2.

4.1.4 Extraction des objets mobiles

Dans cette section, l'objectif est de regrouper les points mobiles appartenant à un même objet, ainsi il sera possible de distinguer plusieurs objets dans une même scène. Pour cela, nous utilisons un étiquetage des composantes connexes [Suzuki *et al.*, 2003]. Cette méthode consiste à affecter une étiquette (un nombre) différente à tous les points détectés comme mobiles. Dans la figure 4.4a), les cases colorées correspondent à des points détectés comme mobiles. Ainsi dans la figure 4.4b), l'algorithme assigne un nombre différent à chaque point mobile. Ensuite, la technique consiste à regarder pour chaque point p , si un de ses voisins a un nombre plus petit. Si ce cas est avéré, l'étiquette du voisin le plus petit est attribuée au point p . Cette opération est montrée à la figure 4.4c). Chaque point prend à l'étape $t+2$ le nombre le plus petit parmi ses voisins de l'étape $t+1$. Cette commande est répétée jusqu'à ce qu'il n'y ait plus de changement dans l'affectation des nombres aux points de l'image. La figure 4.4d) est la dernière étape de l'algorithme, car les étiquettes des points ne changent plus. Ainsi, tous les points ayant un même nombre seront regroupés comme étant un même objet mobile. Dans la figure 4.4d), l'algorithme a détecté deux objets différents, un objet avec les points portant le numéro 1 et un autre objet avec les points ayant le numéro 7.

Cet algorithme pose la question du nombre de voisin à prendre en compte. Est-ce qu'il faut se limiter aux points juxtaposés de p ou alors déterminer un voisinage plus grand? Un des problèmes, à ne considérer que les voisins juxtaposés du point (c'est-à-dire avec un voisinage de

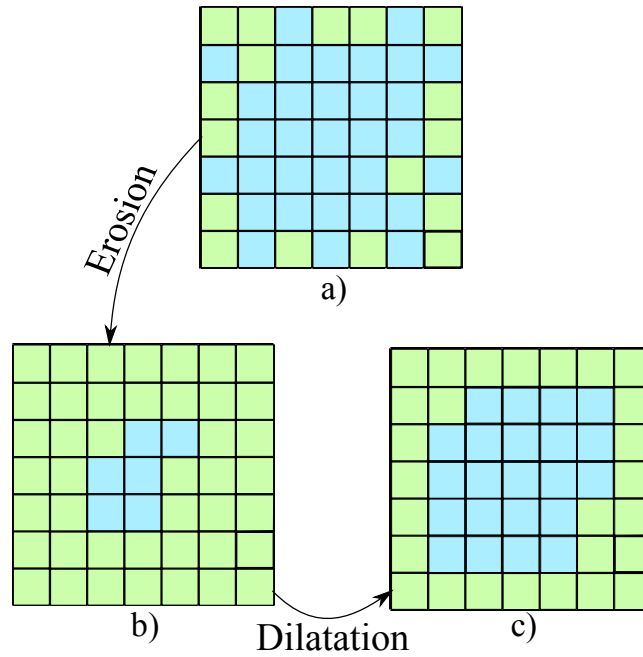


FIGURE 4.3 – a) Les cases bleues représentent des points mobiles et les cases vertes les points appartenant au fond, b) le filtre érosion est appliqué, c) le filtre dilatation est appliqué.

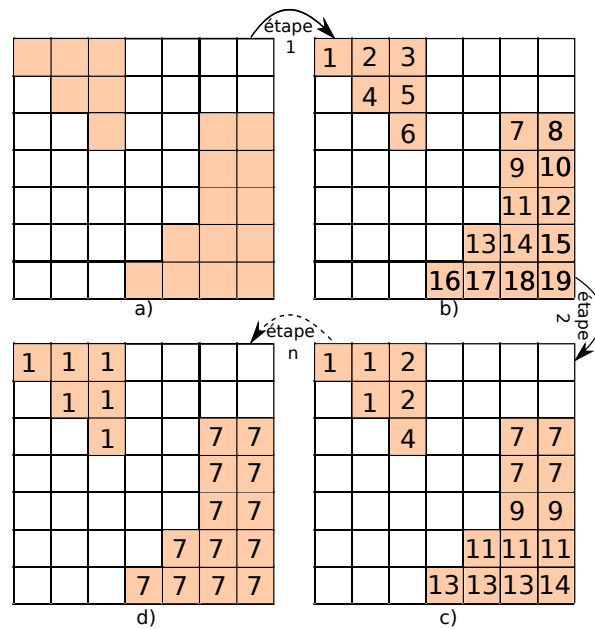


FIGURE 4.4 – a) Les cases colorées représentent des points mobiles, b) l'algorithme assigne un nombre aux points mobiles, c) les points avec un nombre plus grand que ses voisins sont modifiés, d) le résultat final est atteint.

taille 1), est qu'il est possible que certaines parties ne soient pas détectées comme appartenant au même objet. Par conséquent, la jambe ou le bras peut être identifié comme un objet et le reste du corps comme un autre objet. La seconde solution est de prendre un voisinage de taille plus importante. L'algorithme avec un voisinage de taille 2 considérera les points positionnés en $p-2$, $p-1$, $p+1$ et $p+2$. La contrepartie est alors le risque d'intégrer des points n'appartenant pas à l'objet. Pour illustrer ces deux méthodes, nous pouvons donner l'exemple suivant. La figure 4.4 montre le cas où nous considérons seulement les points juxtaposés. Deux objets sont alors détectés en figure 4.4d). Un voisinage de taille 3 modifierait le résultat, les deux blocs de points mobiles seraient rassemblés en un seul objet.

Après une analyse qualitative, en variant la taille du voisinage, un voisinage de 3 est sélectionné pour notre algorithme.

4.1.5 Détection de la personne

Après l'étape précédente, l'algorithme est capable de fournir tous les groupes d'objets mobiles visibles dans la scène. Nous faisons l'hypothèse qu'il n'y a qu'une personne dans la scène. Donc parmi les objets mobiles se trouvent la personne à suivre et des objets qui ont été déplacés (tels qu'une chaise, une porte). Le groupe correspondant à la personne à suivre est récupéré en prenant le groupe ayant le nombre de points mobiles le plus important.

Comme nous l'avons précisé précédemment, le fond ne cesse d'être appris ce qui permet aux objets déplacés de revenir dans le fond au bout d'un certain temps. Par contre, le même effet se produit si une personne reste inactive pendant un certain temps. Pour éviter que la personne n'intègre le fond, l'algorithme d'apprentissage du fond ne prend pas en compte les pixels détectés comme mobiles et appartenant à la personne suivie.

4.2 Suivre la personne

L'objectif de cette étape est de décrire la méthode utilisée dans l'algorithme pour suivre la personne au cours du temps. Comme précisé au chapitre précédent, la méthode pour suivre une personne découle du choix pris pour la représenter. Nous introduisons dans cette section une partie concernant le choix de la représentation, pour ensuite décrire la méthode pour obtenir cette représentation au cours du temps.

4.2.1 Représentation de la personne

Comme nous l'avons présenté dans le chapitre précédent, plusieurs représentations de la personne sont possibles, par des points, par un squelette, par une ellipse représentant la totalité du corps ou une pour chaque partie du corps, etc. Notre objectif était de trouver une représentation assez légère en temps de calcul pour que le suivi se fasse en temps réel.

Les étapes précédentes de détection de la personne nous permettent d'obtenir une série de points 3D la composant. Nous n'avons pas besoin de suivre très précisément les différentes parties du corps, une analyse globale de la personne est suffisante. Nous pouvons réduire les paramètres à un ensemble d'éléments descriptifs de la distribution verticale du corps. Un intérêt majeur de travailler avec une représentation simple est qu'elle sera moins sujette à variations et donc robuste à la diversité des situations.

Le suivi des déplacements du corps dans l'espace peut être une tâche assez complexe, car chaque partie du corps peut faire l'objet d'une analyse à part entière. L'article de Saunders *et al.* [Saunders *et al.*, 1953] traite du sujet des déplacements du corps en mouvement. Selon eux,

l'analyse du corps en mouvement peut être réduite à la simple analyse du centre de masse en mouvement. Plus précisément, les déplacements horizontaux et verticaux du corps, lorsqu'une personne marche, peuvent être analysés en regardant les déplacements du centre de masse. La figure 4.5 et 4.6 représentent les déplacements horizontaux et verticaux du centre de masse quand une personne marche d'après l'article de Saunders *et al.*. Ces déplacements forment des courbes sinusoïdales d'une certaine amplitude. Ce sont ces mécanismes effectués ainsi qui permettent de minimiser la dépense énergétique au cours de la marche. Si les déplacements du centre de masse d'une personne ne respectent pas les formes et l'amplitude habituelle des courbes, la marche est alors dite pathologique. La trajectoire du centre de masse à elle seule permet de fournir de nombreuses informations sur le corps en mouvement.

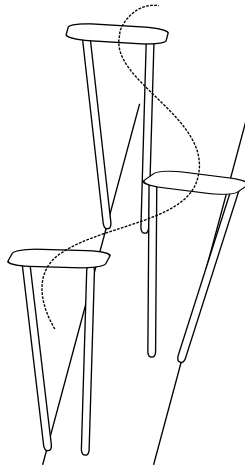


FIGURE 4.5 – Déplacement horizontal théorique du centre de masse.

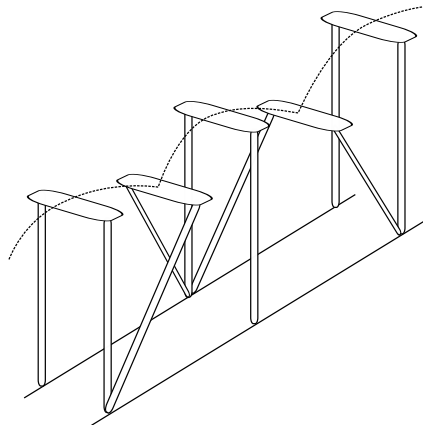


FIGURE 4.6 – Déplacement vertical théorique du centre de masse.

Pour représenter le corps de la personne et le suivre au cours du temps, nous avons choisi de coupler l'analyse de la distribution verticale et de la trajectoire du centre de masse pour représenter le corps durant le cycle de déplacement.

4.2.2 Suivi du centre de masse

Calcul du centre de masse

Au début de cette étape, la personne à suivre est détectée. La personne extraite avec l'algorithme est composée d'un ensemble de points 3D. Dans notre algorithme le centre de masse est calculé comme étant un centre géométrique de la silhouette, il s'agit de l'emplacement moyen, de tous les éléments mobiles, calculé à partir de la formule suivante :

$$Cm(x, y, z) = (\frac{1}{N} \sum_i x_i, \frac{1}{N} \sum_i y_i, \frac{1}{N} \sum_i z_i)$$

où x_i, y_i, z_i sont les coordonnées du i^{eme} points 3D et N est le nombre de points mobiles. La trajectoire du centre de masse d'une personne marchant perpendiculairement à la caméra, obtenue avec notre méthode, est montrée à la figure 4.7. La trajectoire horizontale de la figure 4.7(b) est difficilement interprétable. La trajectoire verticale de la figure 4.7(a) est, en revanche, similaire à la courbe théorique sinusoïdale que Saunders *et al.* ont décrit et qui est représentée à la figure 4.6. De plus, Saunders *et al.* spécifient que l'amplitude bas-haut des déplacements du centre de masse est approximativement 1,08 pouce (soit 4,8 cm), ce qui correspond à peu près à l'amplitude trouvée avec nos courbes.

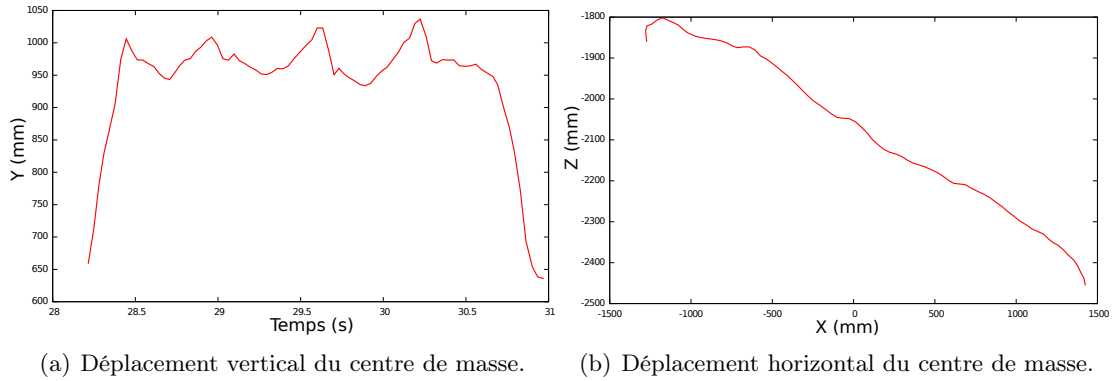


FIGURE 4.7 – Déplacement du centre de masse extrait à partir de notre algorithme pour une personne marchant en ligne droite perpendiculairement à la caméra. La montée et la descente en début et fin de trajectoire correspondent aux entrées et sorties de la personne dans le champ de vision de la caméra.

Description de la trajectoire du centre de masse

L'amplitude de la courbe verticale du centre de masse est accentuée par un artefact lié aux occlusions périodiques faisant varier la répartition du nombre de points visibles appartenant à la personne. La figure 4.8(a) montre que les amplitudes de la courbe du centre de masse sur l'axe vertical sont dépendantes de la variation du nombre de pixels mobiles. Sur cette figure, les maxima locaux de la trajectoire du centre de masse, pour une personne marchant perpendiculairement à la caméra, correspondent aux instants où la personne a moins de pixels mobiles détectés. La raison est qu'à cet instant la personne est sur un pied d'appui et qu'une jambe est cachée par l'autre. Quand la personne est en double appui, les deux jambes sont visibles et il y a donc plus de pixels mobiles détectés dans la partie basse de la silhouette.

Ensuite la trajectoire du centre de masse, d'une personne marchant cette fois en direction de la caméra, a été tracée sur le même graphe que la courbe représentant le nombre de pixels mobiles. Le résultat est montré à la figure 4.8(b). Le nombre de pixels mobiles augmente de plus en plus parce que la personne s'approche de la caméra et occupe donc plus de place dans le champ de vision. La trajectoire du centre de masse s'élève lorsque la personne se rapproche de la caméra. Ceci est dû au fait que, la caméra étant positionnée en hauteur et plongeant vers le bas, la zone supérieure du corps de la personne est de plus en plus visible au fur et à mesure que la personne s'approche de la caméra. La trajectoire du centre de masse sur l'axe vertical forme une trajectoire sinusoïdale mais avec une amplitude plus petite comparée à la situation où la caméra est perpendiculaire à la personne comme à la figure 4.8(a). Dans le cas où la caméra est face à la personne, l'artefact du nombre de points mobiles n'intervient pas. Comme on peut le voir sur la figure 4.8(b), la dépendance entre nombre de points et maxima locaux sur la courbe du centre de masse n'est pas évidente.

Ceci montre que l'artefact, dû aux occlusions de profil, amplifie le mouvement vertical du centre de masse mais n'est pas le seul facteur à l'origine de ce mouvement et, lorsque la caméra est de face, c'est bien le déplacement vertical de la personne qui est observé.

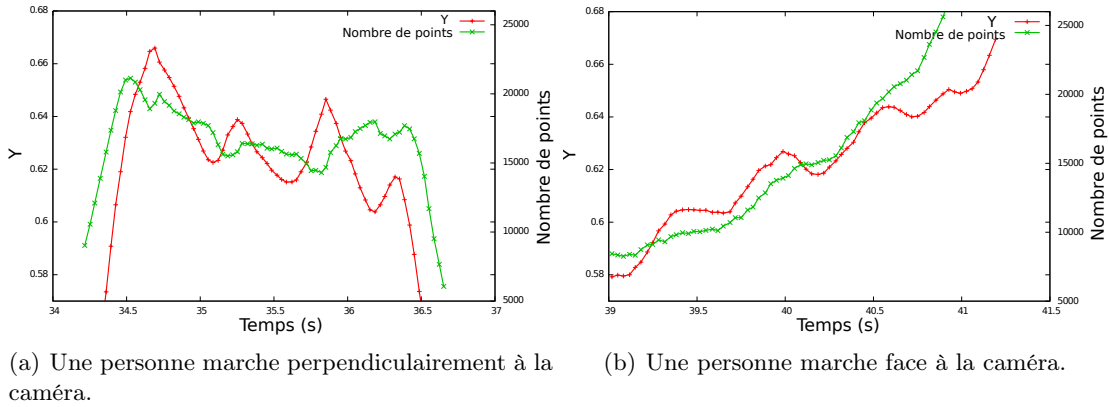


FIGURE 4.8 – Comparaison entre la trajectoire du centre de masse et le nombre de pixels mobiles appartenant à une personne.

Lissage de la trajectoire du centre de masse

Les trajectoires, sur l'axe vertical du centre de masse que nous obtenons, sont filtrées. Deux méthodes pour filtrer la trajectoire ont été implémentées, un filtre passe-bas et un filtre de Kalman.

Notre filtre passe-bas moyenne chaque centre de masse g_t avec les deux centres de masse précédents g_{t-1}, g_{t-2} et les deux centres suivants g_{t+1}, g_{t+2} . Cette méthode permet de filtrer la courbe. Le résultat du lissage de la courbe 4.9(a) avec le filtre passe-bas est montré à la figure 4.9(b).

Le filtre de Kalman [Rabiner, 1989] est un cas particulier du filtre bayésien général, estimant l'état d'un système à partir d'une série d'observations bruitées. Dans notre cas, les observations sont les positions 3D du centre de masse. Le modèle est paramétré par, d'un côté, l'incertitude de la prédiction et, de l'autre, par l'incertitude de l'observation. Le bruit relatif R donné à l'observation, comparé au bruit de la prédiction Q , permet de définir l'inertie du filtre. Nous

utilisons un filtre simple défini par :

$$\begin{aligned}x_{t+1} &= x_t + Q \\g_t &= x_t + R\end{aligned}$$

où g_t représente l'observation à l'instant t (centre de masse calculé sur l'image t) et x_t représente le résultat du filtre. Les bruits Q et R du filtre sont réglés manuellement.

L'inférence est réalisée en deux étapes avec les équations de Kalman.

Prédiction L'étape de prédiction permet d'estimer la position à l'instant t à partir des observations précédentes (des instants 1 à $t - 1$).

$$\begin{aligned}x_{t|t-1} &= Fx_{t-1|t-1} \\P_{t|t-1} &= FP_{t-1|t-1}F^T + Q\end{aligned}$$

Dans notre cas, la fonction de transition F est égale à la matrice identité. $x_{t|t-1}$ représente l'état prédit sans prendre encore en compte l'observation à l'instant t et $P_{t|t-1}$ représente la covariance de la prédiction.

Correction L'étape de correction consiste à mettre à jour la prédiction en prenant en compte l'observation à l'instant t .

$$\begin{aligned}y_t &= g_t - Hx_{t|t-1} \\S_t &= HP_{t|t-1}H^T + R \\K_t &= P_{t|t-1}H^TS_t^{-1} \\x_{t|t} &= x_{t|t-1} + K_ty_t \\P_{t|t} &= (I - K_tH)P_{t|t-1}.\end{aligned}$$

Dans notre cas, la fonction d'observation H est égale à la matrice identité. y_t l'innovation, S_t la covariance de l'innovation et K_t le gain de Kalman sont des variables intermédiaires utilisées pour calculer l'état $x_{t|t}$ et la covariance $P_{t|t}$ a posteriori en intégrant l'observation g_t .

Le résultat du lissage de la courbe 4.9(a) avec le filtre de Kalman est montré à la figure 4.9(c). La comparaison visuelle des deux filtres montre que le filtre de Kalman est moins sensible aux valeurs extrêmes.

Paramètres dynamiques

En plus de la position du centre géométrique, nous utilisons les paramètres représentatifs de son déplacement à savoir sa vitesse horizontale et verticale.

4.2.3 Suivi de la distribution verticale

Pour compléter la représentation de la personne, nous ajoutons deux autres paramètres représentatifs de la distribution verticale qui sont l'écart type et le maximum des coordonnées verticales des points mobiles.

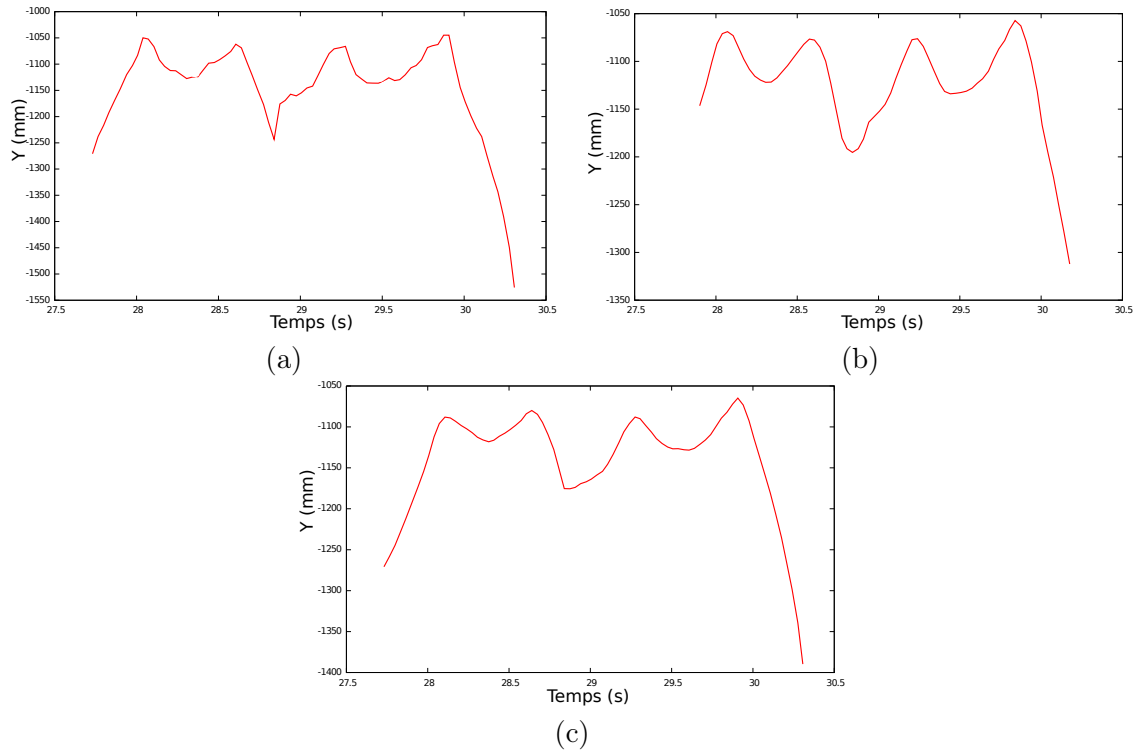


FIGURE 4.9 – a) Courbe non lissée, b) courbe lissée avec le filtre passe-bas, c) courbe lissée avec le filtre de Kalman.

L'écart type des coordonnées des points 3D mobiles, composant le corps de la personne est calculé sur l'axe y par la formule suivante :

$$\sigma_y = \sqrt{\frac{\sum_{i=1}^N [y_i - Cm(y)]^2}{N}}$$

avec N le nombre de points composant le corps de la personne et $Cm(y)$ la coordonnée sur l'axe y du centre de masse. Un exemple de variation de la distribution verticale pour une personne qui marche est illustré à la figure 4.10.

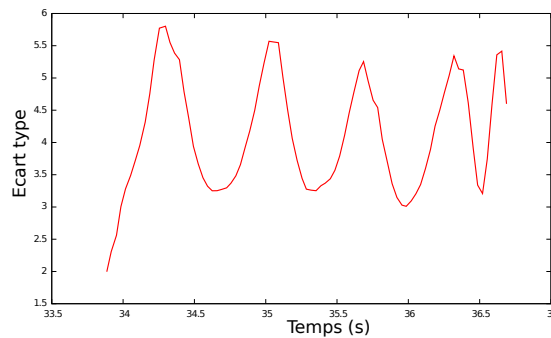
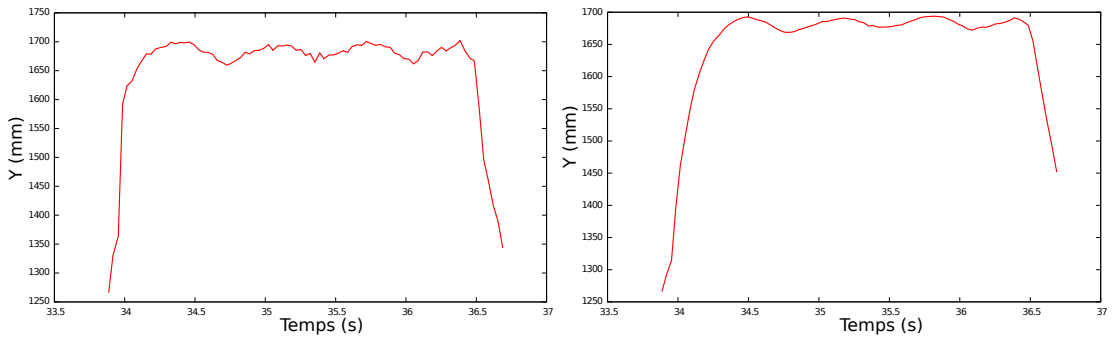


FIGURE 4.10 – Trajectoire de l'écart type vertical pour une personne qui marche.

Le point maximum des points mobiles sur l'axe vertical y est calculé par la formule suivante :

$$Y_{max} = \max \sum_{i=1}^N y_i$$

avec N le nombre de points composant le corps de la personne et y_i la coordonnée sur l'axe y du point mobile i . Tout comme pour le centre de masse, la trajectoire du point maximum est lissée par un filtre de Kalman. La figure 4.11(a) représente la trajectoire du point maximum lorsqu'une personne marche et la figure 4.11(b) correspond à sa trajectoire filtrée. Nous pouvons voir que l'amplitude des oscillations du point maximum est moins grande que celle des oscillations du centre de masse représenté à la figure 4.7(a).



(a) Trajectoire du point maximum sur l'axe vertical. (b) Trajectoire du point maximum lissée sur l'axe vertical.

FIGURE 4.11 – Comparaison entre la trajectoire du point maximum filtrée et non filtrée pour une personne qui marche.

4.3 Conclusion

L'objectif pour détecter et suivre la personne était d'utiliser des algorithmes simples. Ainsi, notre méthode nous permet un suivi de la personne en temps réel.

Nous avons opté pour une représentation de la personne à travers 5 paramètres faciles à extraire qui sont :

- la position du centre de masse ;
- la vitesse verticale du centre de masse ;
- la vitesse horizontale du centre de masse ;
- la dispersion verticale de la silhouette ;
- le point maximum de la silhouette.

Les traitements de plus haut niveau qui sont décrits par la suite (l'analyse des paramètres de la marche et la détection de l'activité d'une personne) se basent sur l'analyse d'un ou de l'ensemble de ces paramètres.

Chapitre 5

Mesurer à domicile les paramètres de la marche

Sommaire

5.1	Définition et description de la marche	58
5.1.1	Définition de la marche	58
5.1.2	Description du cycle de la marche	58
5.2	Indicateurs de l'analyse de la marche	59
5.2.1	Paramètres spatio-temporels	59
5.2.2	Paramètres cinématiques	60
5.2.3	Paramètres dynamiques	61
5.3	Sélection des indicateurs	62
5.3.1	Troubles de la marche et risques de chute	64
5.3.2	Troubles de la marche et risques d'hospitalisation	64
5.3.3	Troubles de la marche et troubles cognitifs	64
5.4	Extraction des indicateurs	65
5.4.1	Longueurs de pas	65
5.4.2	Cadence	66
5.4.3	Vitesse de marche	66
5.4.4	Résultat	66
5.5	Conclusion	67

L'être humain effectue en moyenne 5 000 à 15 000 pas par jour soit 2 à 5 millions par an. La marche s'acquiert au travers d'un processus d'apprentissage qui est qualifié de difficile selon Viel [Viel, 2000]. L'apprentissage de la marche prend plusieurs années, pour être par la suite automatique. Le but pour les nouveaux apprenants est d'acquérir les différentes phases composant la marche, de maintenir leur équilibre et de s'adapter aux contraintes de l'environnement extérieur tout en réduisant le coût énergétique. L'objectif dans ce chapitre est de décrire ce qu'est la marche. Nous commençons par décrire la marche dans sa globalité en analysant les différentes phases qui la constituent. Puis une seconde section abordera la façon dont la marche peut être analysée en énumérant différents indicateurs utilisés pour la décrire. Cette section est inspirée des thèses de Gillet en 2004 [Gillet, 2004] et Fusco en 2008 [Fusco, 2008]. Après avoir brossé un panorama des indicateurs existants pour analyser la marche, nous proposons de sélectionner certains paramètres pertinents pour l'étude de la fragilité. Enfin, nous décrivons l'algorithme qui permet d'extraire ces paramètres.

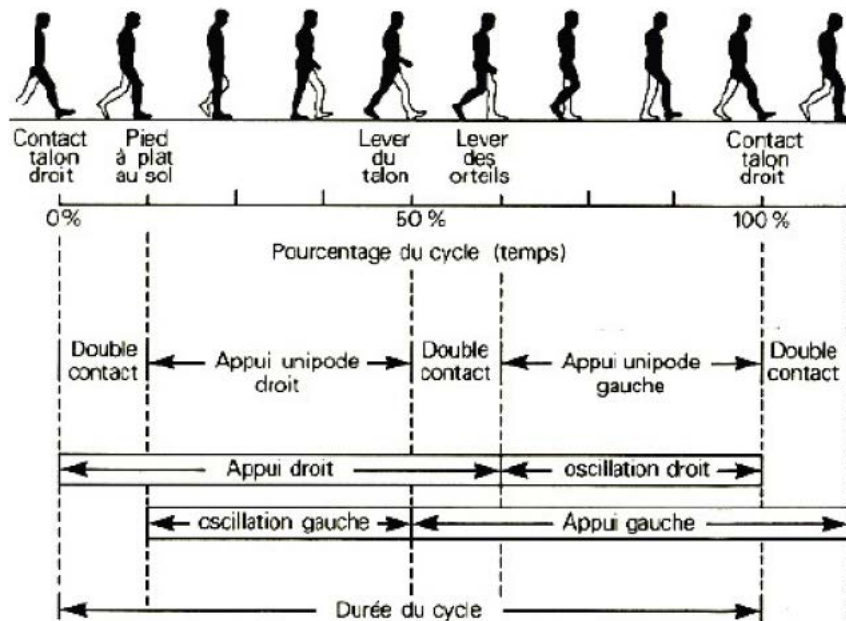


FIGURE 5.1 – Représentation du cycle de la marche selon Viel (2000).

5.1 Définition et description de la marche

5.1.1 Définition de la marche

La marche humaine « normale » est définie par Perry [Perry and Davids, 1992] comme un phénomène complexe impliquant la coordination de mouvements de rotations des segments corporels pour maintenir l'équilibre du corps pendant son déplacement vers l'avant. Chaque être humain a sa propre façon de marcher, sa propre démarche. Malgré tout, la marche peut être découpée en séquences de mouvements cycliques, qui sont identifiables chez tous les humains, pour les deux sexes et âges confondus, et qui apparaissent à chaque pas. La description de la marche, ci-après, concerne la marche « normale », par opposition à la marche pathologique.

5.1.2 Description du cycle de la marche

Un cycle de marche (une foulée) commence par le contact initial d'un pied au sol (par le talon) et se termine par le contact suivant du même pied (par le même talon). Le cycle de la marche est lui-même découpé en phases, permettant ainsi de poser les bases pour analyser la marche d'une personne. Selon le découpage de Viel [Viel, 2000], le cycle de marche est découpé principalement en deux phases, une phase d'appui et une phase d'oscillation comme illustré à la figure 5.1.

La phase d'appui (droite ou gauche) dure 60 % du cycle et est définie comme la phase débutant lorsque le pied (droit ou gauche) est en contact avec le sol. La phase d'oscillation (droite ou gauche) dure 40 % du cycle et est définie comme la phase débutant lorsque la jambe (gauche ou droite), en l'air, se déplace vers l'avant et se termine lorsque le talon reprend contact avec le sol. Un découpage plus fin du cycle peut être donné à l'intérieur des phases d'appui et d'oscillation.

Phase d'appui

Le double appui de réception Lors de cette phase, les deux pieds sont en contact avec le sol. Cette phase commence à 0 % du cycle avec le contact initial du talon sur le sol et se termine lorsque les orteils controlatéraux se soulèvent à 10 % du cycle. Le rôle de cette phase est de diriger le mouvement du corps vers l'avant et vers le côté lorsque le pied controlatéral quitte le sol.

La phase d'appui unilatéral Après la phase de double appui de réception se produit la phase d'appui unilatéral qui dure environ 40 % du cycle. Lors de cette phase, un seul membre inférieur est en appui. La phase commence à 10 % du cycle avec le décollement des orteils controlatéraux et se termine à environ 50 % du cycle avec le contact du talon controlatéral. Cette période permet la progression du corps en avant par le pied en appui.

Le double appui de propulsion Tout comme la phase de double appui de réception, les deux pieds sont en contact avec le sol. Cette phase commence lors du contact controlatéral à environ 50 % du cycle et se finit à 60 % du cycle avec le décollement des orteils du pied en appui. Le but de cette phase est de propulser le corps vers l'avant, grâce au transfert de poids vers la jambe controlatérale.

Phase d'oscillation

Durant cette phase, le membre oscillant ne décolle que très légèrement les orteils (1,29 cm) puis le talon frôle la surface du sol d'environ 1 cm. Habituellement, cette phase n'est pas découpée plus finement. Nous pouvons tout de même noter que, lors de cette période, une phase d'accélération et une phase de décélération peuvent être distinguées. La phase d'accélération commence lorsque le pied décolle du sol et se termine lorsque les deux pieds se croisent. Ensuite une phase de décélération est observée dans laquelle le membre oscillant ralentit dans le but de préparer le prochain contact.

5.2 Indicateurs de l'analyse de la marche

Pour analyser la marche d'une personne, différentes variables peuvent être prises en compte selon le but de l'analyse. Les indicateurs peuvent être regroupés selon trois catégories.

5.2.1 Paramètres spatio-temporels

Tout d'abord, les paramètres spatio-temporels sont couramment étudiés, car ils informent sur les caractéristiques globales de la marche, c'est-à-dire qu'ils informent sur le déroulement de la marche dans le temps et l'espace. Les paramètres spatio-temporels ont la particularité de prendre des valeurs variables selon le sexe et l'âge de la personne.

Principaux paramètres spatiaux

Les paramètres souvent évalués sont la longueur du pas, la largeur du pas, la longueur et largeur d'enjambée et, l'angle du pas. Ces paramètres ont une définition parfois différente selon les articles. Il n'y a pas vraiment de consensus en analyse de la marche. Nous proposons ici des définitions inspirées de ce qui est le plus couramment proposé et nous reprenons également la thèse de Faivre de 2003 qui propose un lexique terminologique [Faivre, 2003].

La longueur du pas (droit pour notre exemple) correspond à la distance entre les deux talons lors du temps du double appui de réception (double appui de réception droit dans notre exemple). La largeur du pas est définie comme la distance entre la ligne de progression (ligne traversant tous les pas effectués par la personne) et le point médian du talon. L'enjambée correspond à la succession de deux pas. La longueur d'enjambée est alors définie comme la somme algébrique des longueurs de deux pas successifs. La largeur d'enjambée quant à elle est déterminée par la somme algébrique de la largeur de deux pas successifs. Ces paramètres sont exprimés en mètres. L'angle du pas, exprimé en degrés, correspond à l'angle formé entre la ligne de progression et le point médian du talon. Tous ces paramètres sont représentés à la figure 5.2.

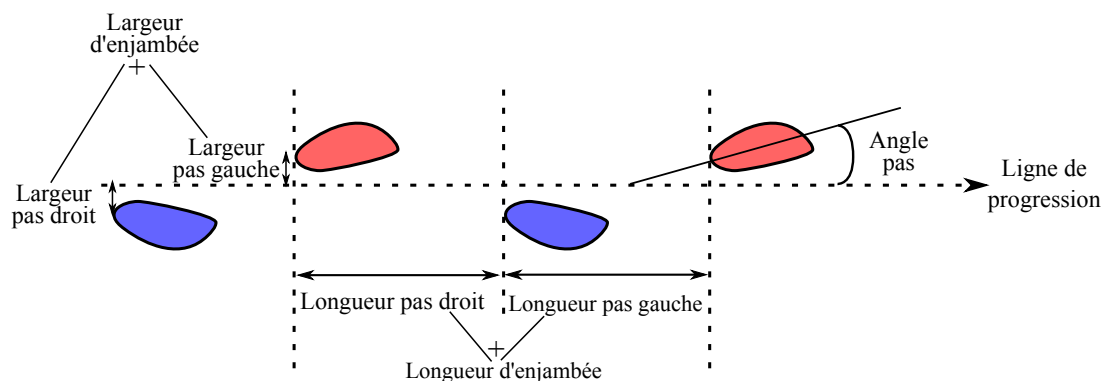


FIGURE 5.2 – Représentation des paramètres spatiaux.

Principaux paramètres temporels

Les paramètres souvent évalués sont la cadence (la fréquence), la vitesse de marche, le temps d'appui bipodal et unipodal (ou temps d'oscillation). Tout comme les paramètres spatiaux, les définitions de chaque terme ne sont pas des définitions « universelles ».

La cadence correspond à la vitesse de marche en fonction du nombre de pas. Il s'agit de la fréquence à laquelle une personne marche. Le terme « fréquence » est d'ailleurs souvent employé à la place de « cadence ». La cadence est calculée comme le nombre de pas par minute ou par seconde. La vitesse de marche correspond à la relation mathématique entre la cadence (enjambées/min) et la longueur d'enjambée (en mètre). La vitesse de marche est exprimée en mètre par seconde. La formule est donc :

$$\text{Vitesse de marche} = (\text{Longueur} \times \text{Cadence}) / 120$$

Bien souvent dans la réalité, la vitesse de marche est obtenue en chronométrant le temps mis par le sujet pour parcourir une distance donnée. Le temps d'appui bipodal est la durée moyenne des temps de double appui lors d'un cycle de la marche (pour rappel le cycle de marche correspond au moment entre le contact initial du talon et le contact suivant du même talon). Le temps d'appui unipodal est la durée moyenne du temps d'appui unilatéral lors d'un cycle de marche droit ou gauche. Ce temps est égal au temps d'oscillation du pied controlatéral. Les temps d'appui bipodal et unipodal sont exprimés en secondes ou en % du cycle de marche.

5.2.2 Paramètres cinématiques

La deuxième catégorie correspond aux paramètres cinématiques qui sont des variables fréquemment étudiées, car elles reflètent le mouvement des membres inférieurs au cours de la marche.

Il s'agit d'analyser l'aspect de la marche. Les paramètres correspondent aux angles permettant d'analyser la flexion de la cheville, la flexion du genou et la flexion de la hanche. La flexion de la cheville est analysée en observant l'angle formé entre le pied et le tibia (os du membre inférieur). La flexion du genou correspond à l'angle entre le tibia et le fémur (os de la cuisse). Et la flexion de la hanche représente l'angle formé entre le fémur et le bassin. Ces angles sont représentés à la figure 5.3.

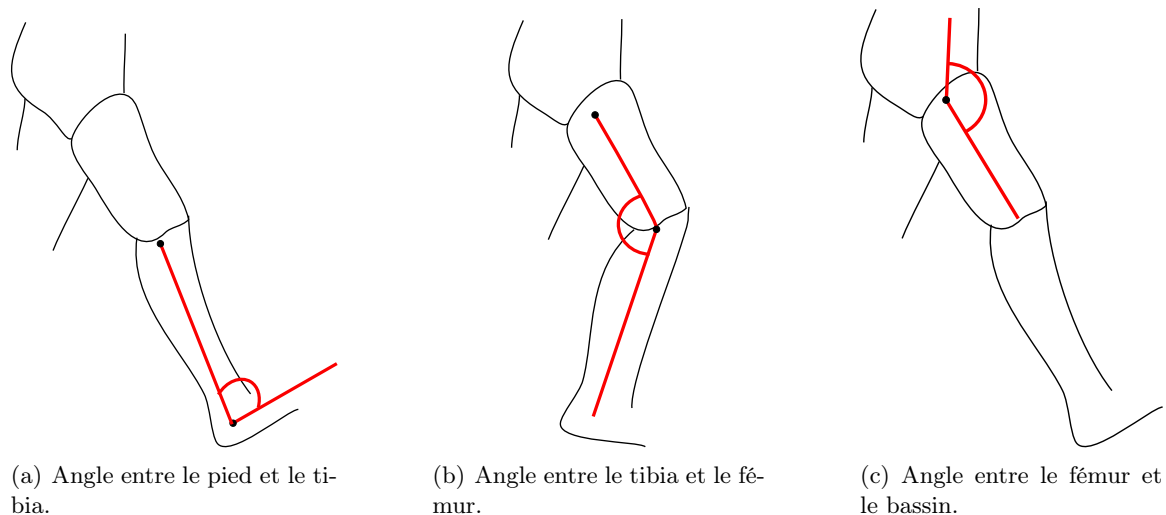


FIGURE 5.3 – Représentation des paramètres cinématiques.

5.2.3 Paramètres dynamiques

L'analyse dynamique constitue l'étude des forces et moments internes qui engendrent le mouvement. L'utilisation de la technique de l'électromyographie décrite dans la section 1.2.1 permet cette analyse. Mais celle-ci est invasive, elle requiert de poser des capteurs sur les muscles. Pour faciliter l'analyse dynamique, il est courant d'étudier les forces et les moments externes à travers l'analyse de l'interaction du pied avec le sol. Ainsi, les forces de réaction du sol et les pressions plantaires permettent une évaluation du pied.

Les forces de réaction du sol

Les forces de réaction du sol sont mesurées lors de la réception et la propulsion des pieds durant la marche. La force de réaction correspond à un vecteur se divisant selon trois axes d'après Kirtley [Kirtley, 2006], un axe vertical, antéro-postérieur et média-latéral.

Composante verticale La composante verticale correspond à la force décrite à la figure 5.4(a). Il s'agit de la composante où les forces ont le plus d'amplitude. Lors d'un cycle de la marche, la courbe de cette composante évolue en forme de « M » comme représentée en bleue à la figure 5.5. Les deux pics de cette courbe représentent les phases de réception et de propulsion. Lors de la phase d'oscillation, le pied n'est plus en contact avec le sol donc la force de réaction est nulle.

Composante antéro-postérieure La composante antéro-postérieure correspond à la force décrite à la figure 5.4(b). Deux phases sont identifiables sur la courbe formée par cette composante

lors du cycle de la marche comme montré en vert à la figure 5.5. Lors de la phase de réception, la courbe est négative, le sujet freine, donc la force est exercée vers l'arrière. La deuxième phase, la phase de propulsion, est le moment où la courbe est positive, le sujet propulse son corps vers l'avant donc la force est exercée vers l'avant.

Composante média-latérale La composante média-latérale, illustrée à la figure 5.4(c), correspond aux forces qui ont le moins d'amplitude. Toutefois il est possible de constater deux phases dans la forme de la courbe des forces, comme le montre en rouge la figure 5.5. Une phase de réception très courte où la force est orientée vers l'extérieur, vers l'axe latéral. Et une phase d'appui qui dure plus longtemps et où la force est orientée vers le centre, vers l'axe médian.

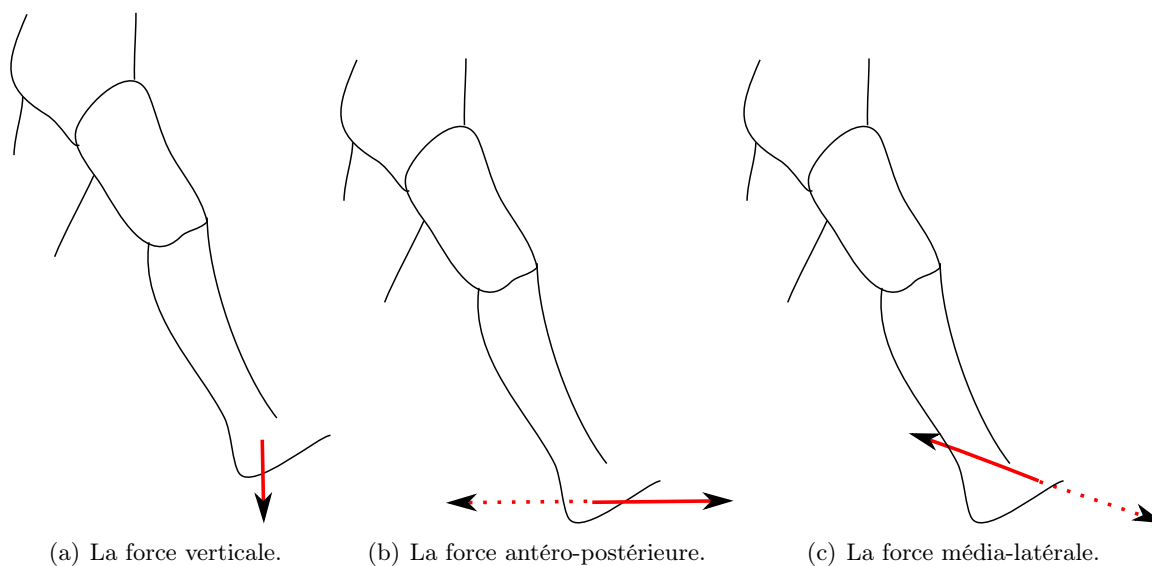


FIGURE 5.4 – Représentation des forces exercées lorsqu'une personne marche.

Pressions plantaires

Les pressions plantaires correspondent aux points d'appui sur l'ensemble de la plante du pied lors du contact avec le sol. Les points d'appuis successifs lorsque le pied est en contact avec le sol sont représentés à la figure 5.6 et inspirés selon Allard en 1996 [Allard and Blanchi, 1996]. Tout d'abord les pressions débutent au niveau de l'extérieur du talon lorsque la personne pose son talon au sol. Ensuite, la trajectoire se poursuit sur le bord externe du pied, pour se propager vers toutes les têtes métatarsiennes et aboutir au niveau de la phalange distale de l'orteil.

5.3 Sélection des indicateurs

Lorsqu'une personne vieillit, sa marche est modifiée. Plusieurs changements dans la marche des personnes âgées peuvent être cités, notamment une diminution de la vitesse de la marche [Guimaraes and Isaacs, 1980], une réduction des longueurs de pas ([Guimaraes and Isaacs, 1980], [Winter et al., 1990]), une augmentation du temps de double appui [Winter et al., 1990] et une plus grande variabilité de la marche [Guimaraes and Isaacs, 1980]. Ces modifications de la marche sont considérées comme « normales », car dues à l'avancée dans l'âge. Ce qui nous intéresse dans

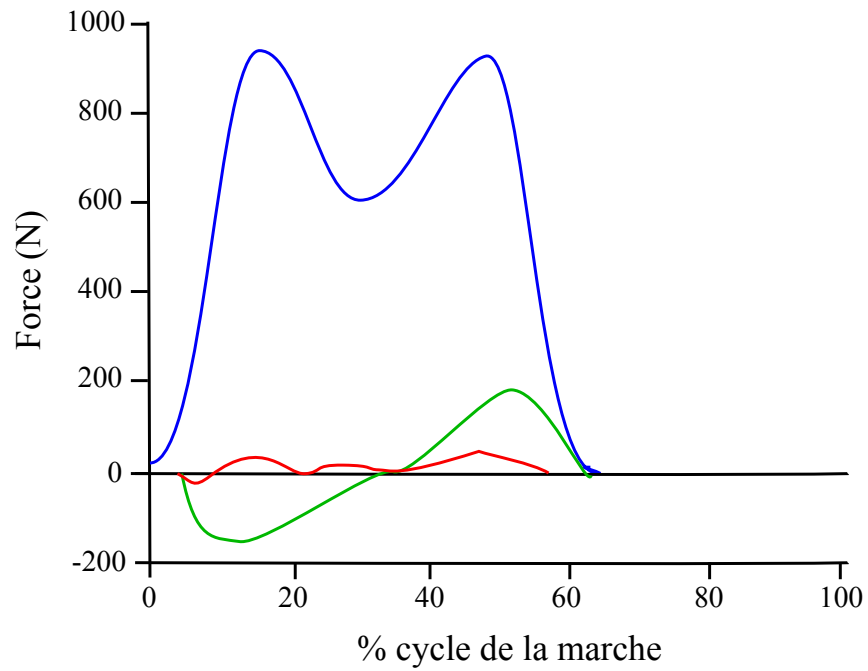


FIGURE 5.5 – Force de réaction du sol selon l'axe vertical (bleu), antéro-postérieur (vert) et média-latéral (rouge) lors d'un cycle de marche selon Kirtley.

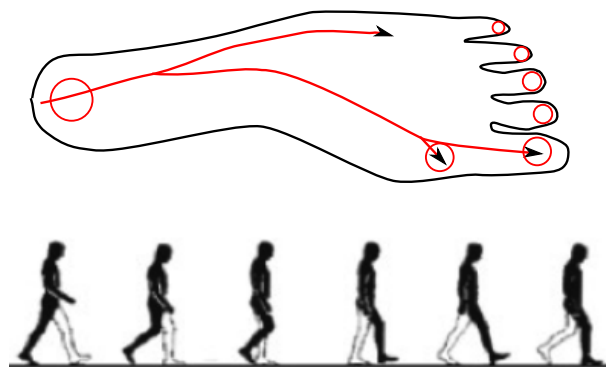


FIGURE 5.6 – Trajectoire des pressions plantaires lors de la marche.

cette partie, concerne les changements « anormaux » de la marche, pouvant être des indicateurs d'une augmentation de la fragilité des personnes.

Dans cette section, nous présentons des articles mettant en évidence une corrélation entre une détérioration de la santé de la personne et une évolution des indicateurs de sa marche. L'analyse de ces articles nous permettra de déterminer quels sont les indicateurs pouvant informer sur l'état d'une personne indépendamment de l'âge.

5.3.1 Troubles de la marche et risques de chute

En 1980, Guimaraes *et al.* [Guimaraes and Isaacs, 1980] ont réalisé une étude sur l'analyse de la marche en comparant deux groupes, un groupe de 30 personnes hospitalisées pour cause de chute (sans blessure particulière) et un groupe de 22 personnes de même âge se trouvant à l'hôpital mais n'ayant pas subi de chute récente. L'étude a révélé que le groupe hospitalisé, pour cause de chute, avait une vitesse de marche réduite, de plus petite longueur et largeur de pas et une augmentation dans la variabilité des fréquences de pas.

En 2001, Hausdorff *et al.* [Hausdorff *et al.*, 2001] reprennent l'idée d'une dégradation des paramètres chez les personnes ayant chuté et testent l'hypothèse que la prévention des chutes est possible à partir de ces paramètres. Après un an d'expérimentation, ils ont montré que les variabilités de temps des foulées et les variabilités de temps d'oscillation sont associées à une augmentation du risque de chute dans le futur.

Auvinet *et al.* [Auvinet *et al.*, 2003] ont mis en place en 2003 une expérimentation où 53 sujets, dont 20 étaient considérés comme le groupe des « chuteurs », devaient marcher sur un parcours de 40 mètres. Ils ont alors pu montrer que la vitesse, la fréquence et les longueurs de pas avaient des valeurs réduites pour le groupe des « chuteurs ». L'irrégularité des longueurs de pas était considérée, dans leur étude, comme une variable fiable pour prévenir les chutes.

5.3.2 Troubles de la marche et risques d'hospitalisation

D'autres articles montrent l'intérêt d'étudier les paramètres spatio-temporels pour leur corrélation avec le risque d'hospitalisation. En effet, l'analyse de la vitesse de la marche serait un bon indicateur pour prédire un risque d'hospitalisation et un déclin de la santé selon Studenski *et al.* [Studenski *et al.*, 2003]. Dans cet article, les auteurs montrent que 41 % des marcheurs lents (<0.6 m/s) sont hospitalisés au moins une fois par an, tandis que seulement 26 % des marcheurs intermédiaires (0.6 - 1.0 m/s) et 11 % des marcheurs rapides (>1.0 m/s) sont hospitalisés.

5.3.3 Troubles de la marche et troubles cognitifs

Les troubles cognitifs les plus répandus chez les personnes âgées sont regroupés sous le terme de démence. La démence est une détérioration des fonctions cognitives, provoquant un changement dans la vie de la personne et une perte d'autonomie. Les démences couramment citées sont la maladie d'Alzheimer et la maladie de Parkinson.

Les désordres liés à la marche sont plus répandus chez les personnes âgées ayant des démences que chez les sujets « normaux ». Iersel *et al.* [Van Iersel *et al.*, 2004] ont comparé ces deux groupes de sujets et en ont déduit que les personnes avec des démences ont une vitesse de marche plus lente, des longueurs de pas plus petites et un temps de double appui plus grand que les sujets âgés sains. La vitesse de marche plus lente chez les personnes avec des démences est mise en évidence également dans l'article de Bramell-Risberg *et al.* [Bramell-Risberg *et al.*, 2005].

Plusieurs études ont également déterminé qu'un dysfonctionnement mesuré par une vitesse de marche lente peut être un indicateur de déficits cognitifs précoces ([Holtzer *et al.*, 2006],

[Kuo *et al.*, 2007], [Waite *et al.*, 2005]). Les déficits cognitifs sont définis comme une dégradation d'un ou des processus cognitifs mais sans provoquer de perte d'autonomie de la personne. Ce trouble peut éventuellement subvenir plusieurs années avant l'apparition des démences.

En sachant que les personnes âgées atteintes de démence ou présentant un déficit cognitif précoce ont une marche altérée, en comparaison aux personnes âgées sans démence et sans déficit, d'autres chercheurs se sont intéressés à l'analyse des indicateurs de la marche pouvant permettre de prévenir une éventuelle apparition de démence ou déficit cognitif dans les mois ou années qui suivent. Bien que le principal signe de démence soit un déclin cognitif, plusieurs articles montrent que des désordres moteurs sont présents au premier stade de la maladie. Selon Camicioli *et al.* [Camicioli *et al.*, 1998], une marche lente pourrait précéder des déficiences cognitives. Leur expérience consistait à évaluer 85 personnes âgées saines durant 3 ans. L'objectif était de déterminer les différences motrices entre les personnes sans problème cognitif et ceux développant des détériorations cognitives, durant 3 ans. Ils remarquèrent alors que les 18 sujets, développant des troubles cognitifs, avaient un ralentissement moteur (constat à travers le temps mis pour parcourir une certaine distance).

Ces différentes études indiquent que l'analyse de la vitesse pourrait améliorer le diagnostic des problèmes cognitifs, en sachant particulièrement que 50 % des personnes affectées par la maladie d'alzheimer sont diagnostiquées et seulement 30 % au premier stade de la maladie ([Ferri *et al.*, 2006], [Rimmer *et al.*, 2005]). De plus, les sujets avec des démences chutent deux ou trois fois plus que les sujets sans déclin cognitif [Shaw, 2002].

5.4 Extraction des indicateurs

L'analyse des paramètres spatio-temporels s'avère être une approche intéressante pour évaluer l'état général d'une personne. Les paramètres auxquels nous allons nous intéresser sont la vitesse de marche, les longueurs de pas et la cadence. Ces paramètres sont extraits à partir des déplacements verticaux du centre de masse de la personne. Plus précisément, les paramètres sont extraits à partir de la courbe lissée (avec le filtre passe-bas ou le filtre de Kalman).

5.4.1 Longueurs de pas

Comme nous l'avons déjà défini à la section 5.2.1, la longueur de pas, dans la littérature, correspond à la distance séparant les deux talons lors du double appui de réception. Notre méthode, pour récupérer la longueur de pas, consiste à chercher la position du pied d'appui. Le centre de masse est au plus haut au moment où il se trouve à la vertical du pied d'appui. La position du pied d'appui peut donc être estimée par la projection sur le sol du centre de masse au moment du point max sur la courbe, comme illustré par la figure 5.7. Dans notre cas, la longueur de pas est calculée comme la distance entre la projection de deux maxima locaux consécutifs de la trajectoire du centre de masse sur l'axe vertical. La distance est obtenue en calculant la distance euclidienne entre les deux points (A et B) en trois dimensions :

$$d_{a,b} = \sqrt{(x_b - x_a)^2 + (y_b - y_a)^2 + (z_b - z_a)^2}$$

L'extraction de la première longueur de pas est montrée à la figure 5.8. Nous exprimons les longueurs de pas en centimètres.

5.4.2 Cadence

Nous calculons la cadence sur une séquence et non pas pour chaque pas. Pour calculer la cadence de l'ensemble des pas d'une séquence, nous appliquons la formule suivante :

$$Cadence = \frac{N}{(\sum_{i=1}^N duree(i))}$$

avec N le nombre de pas réalisés au cours de la séquence et $duree(i)$ la durée du pas i . La cadence est exprimée ici en nombre de pas par seconde.

La durée d'un pas est définie comme étant le temps entre le point de contact avec le sol d'un pied et le point de contact de l'autre pied (contralatéral). Pour obtenir la durée du pas, nous utilisons la même méthode que pour extraire les longueurs de pas. Le temps entre la projection de deux maxima locaux consécutifs sur la trajectoire du centre de masse sur l'axe vertical constitue la durée d'un pas. La durée d'un pas est exprimée en secondes. L'extraction de la durée du premier pas est montrée à la figure 5.8.

5.4.3 Vitesse de marche

Les vitesses de marche moyenne et instantanée peuvent être calculées de la même façon. Dans la section 5.2.1, nous avons précisé que la vitesse de marche était souvent calculée en chronométrant le temps mis par le sujet pour parcourir une distance donnée. Nous reprenons cette idée pour extraire la vitesse moyenne. La vitesse de marche est calculée par la formule suivante :

$$Vitesse = \frac{\sum_{i=1}^N longueur(i)}{\sum_{i=1}^N duree(i)}$$

avec $longueur(i)$ et $duree(i)$ la longueur de pas (ici, exprimée en mètres) et la durée du pas i . Nous obtenons donc la vitesse de marche moyenne exprimée en mètre par seconde. L'extraction de la vitesse de la marche est montrée à la figure 5.8.

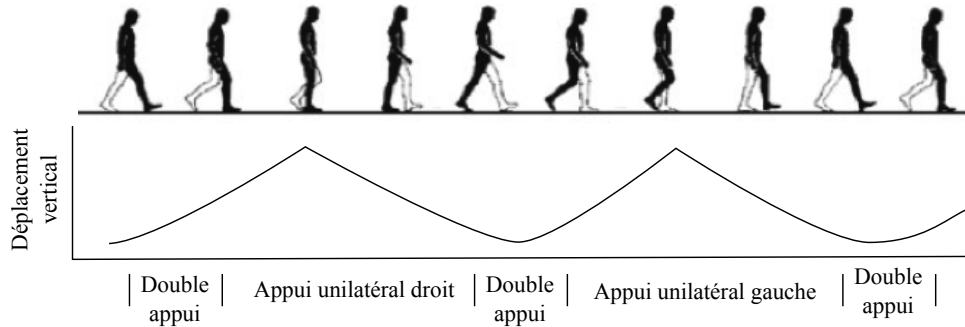


FIGURE 5.7 – Correspondance entre le cycle de la marche et la trajectoire du centre de masse sur l'axe vertical.

5.4.4 Résultat

Lorsqu'une personne marche en ligne droite perpendiculairement à la caméra, nous obtenons entre 2 et 4 pas d'environ 50 à 60 cm selon les personnes. Les derniers pas de chaque séquence (une séquence correspond au moment où la personne traverse le champ de vision de la caméra

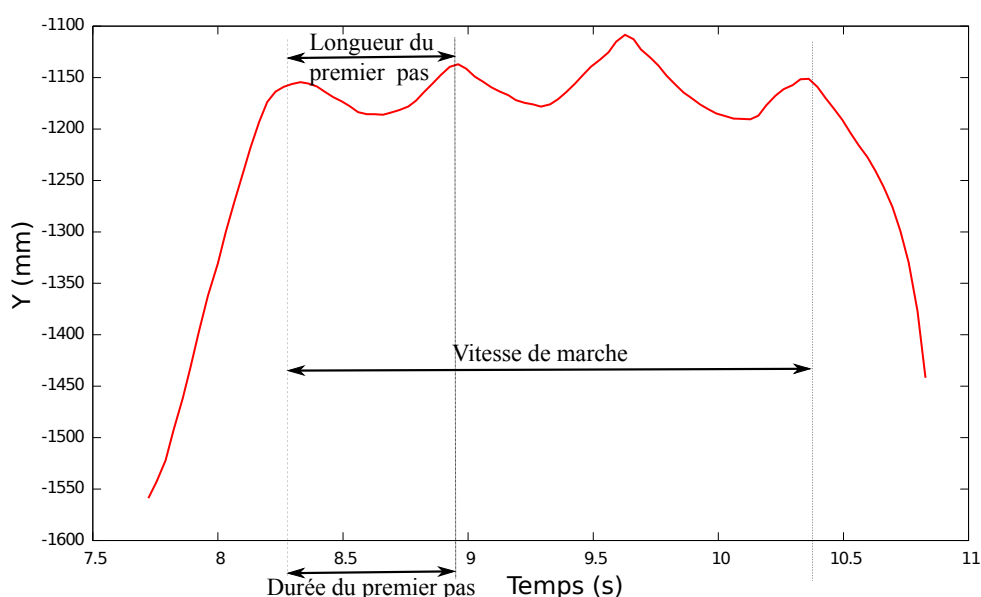


FIGURE 5.8 – Extraction des paramètres de la marche à partir du centre de masse.

en entrant d'un côté et en sortant de l'autre) ne sont pas pris en compte. Les derniers pas correspondent aux moments où la personne sort du champ de vision de la caméra. Une fausse détection du pas peut se produire à proximité des bords de l'image. C'est pourquoi nous préférons enlever le dernier pas de chaque séquence lorsque la personne sort du champ de vision de la caméra.

5.5 Conclusion

Ce chapitre nous a permis de poser les bases et les notions à connaître pour comprendre les différentes phases de la marche. Il liste, également, les différents paramètres existants pour analyser la marche de la personne. Les paramètres spatio-temporels représentent l'analyse des longueurs, cadences et vitesses de marche. Les paramètres cinématiques correspondent à l'analyse des angles formés par le pied, le genou et la hanche. Et les paramètres dynamiques étudient les forces de réaction au sol et les pressions plantaires. Les systèmes analysant la marche des personnes sont fondés sur l'analyse d'un ou de plusieurs de ces paramètres. Par exemple, les tapis actimétriques fournissent les paramètres spatio-temporels et les paramètres dynamiques, les accéléromètres, quant à eux, calculent les paramètres cinématiques. Nous avons également présenté, dans ce chapitre, différentes études mettant en évidence des liens entre troubles de la marche et dégradation de l'état de fragilité de la personne. D'après ces articles, l'analyse de la marche peut permettre de prévenir les chutes, les risques d'hospitalisation et les troubles cognitifs. La plupart des études concordent sur le fait que la longueur de pas, la cadence et la vitesse de marche sont des paramètres fiables pour étudier l'évolution de l'état de santé de la personne. Ainsi, nous avons présenté l'algorithme d'extraction de ces trois paramètres, à partir du centre de masse et de la détection des maxima locaux sur l'axe vertical. L'évaluation de la précision des paramètres extraits est présentée dans la prochaine partie.

Chapitre 6

Reconnaitre l'activité humaine en temps réel par l'analyse d'images de profondeur

Sommaire

6.1	Approches pour détecter l'activité	70
6.1.1	Approches non paramétriques	70
6.1.2	Approches volumétriques	70
6.1.3	Approches paramétriques	71
6.2	Positionnement	71
6.3	Proposition d'un modèle de l'activité	73
6.3.1	Modèle avec fonction d'observation seule	74
6.3.2	Utilisation d'un MMC unique	75
6.3.3	Utilisation de MMC combinés	79
6.4	Conclusion	81

De nombreuses méthodes ont été développées dans le domaine de la reconnaissance d'activité. Les méthodes de reconnaissance d'activité peuvent être regroupées en trois catégories : les méthodes non paramétriques, les méthodes volumétriques et les méthodes paramétriques. En annexe A, nous présentons en particulier une méthode paramétrique, les modèles de Markov cachés (MMC), sur laquelle nous nous sommes basés pour construire nos modèles de détection d'activité. L'un des intérêts des MMC est de posséder de bonnes capacités de généralisation et donc d'être adaptée à nos contraintes. En effet, les données disponibles pour l'apprentissage sont nécessairement limitées et partiellement représentatives de la diversité des situations pouvant être rencontrées en situation réelle.

Nous cherchons à déterminer l'activité d'une personne étant données les mesures que nous pouvons effectuer avec une caméra de profondeur. Cette activité est un état caché, que nous cherchons à déterminer à partir d'une séquence d'observations extraites des images de profondeur de la caméra RGB-D. Ces observations ont été choisies pour leur robustesse et la simplicité avec laquelle elles sont extraites des images de profondeur. Ces indicateurs sont cités au chapitre 4. Il s'agit du centre de masse de la personne, de ses vitesses verticale et horizontale, de la dispersion verticale du corps et du point le plus haut de la silhouette. Nous avons construit plusieurs modèles. Certains sont constitués d'un MMC unique, d'autres plus complexes requièrent plusieurs MMC, un par activité. Nous faisons varier le nombre d'observations et le nombre d'états afin de pouvoir,

ensuite, les comparer pour déterminer le modèle étant le plus pertinent. Cette comparaison est faite dans le chapitre 8.

6.1 Approches pour détecter l'activité

Plusieurs études classifiant les méthodes de détection d'activités à partir d'images vidéo ont été réalisées ces dernières années. Les classifications réalisées sont différentes d'un article à l'autre car la reconnaissance d'activités est un large sujet. Ke *et al.* [Ke *et al.*, 2013] propose ainsi un découpage de la problématique sur trois niveaux, faisant apparaître les grandes étapes de l'analyse d'images, identifiant plusieurs catégories de problèmes en fonction des objectifs et enfin définissant des types d'applications. Les méthodes d'analyse d'images sont en effet liées au dispositif technique et à son environnement. Il faut par exemple prendre en compte les capacités de la caméra choisie, comme le type de capteur, l'ouverture de son champ de vision, la mobilité ou non de la caméra. La classification peut également être faite en fonction des méthodes d'analyse utilisées qui sont très nombreuses et peuvent s'étudier selon les différents niveaux de traitement de l'image et du flux vidéo. Les méthodes d'analyse dépendent en particulier de la complexité et du type des activités à reconnaître. Les méthodes pour détecter qu'une personne lève les bras et se fait un thé ne sont, par exemple, pas les mêmes. Les moyens à mettre en œuvre dépendent enfin du contexte applicatif. Par exemple, la reconnaissance d'activités anormales ne requiert pas les mêmes méthodes si on s'intéresse aux activités dangereuses d'une personne dans le métro ou aux chutes des personnes âgées à leur domicile. Le regroupement des méthodes de reconnaissance d'activités peut donc se faire de nombreuses façons et il est par conséquent difficile d'être exhaustif. Ainsi, nous retrouvons dans la littérature des études se focalisant sur les méthodes appliquées à la surveillance [Valera and Velastin, 2005], sur la détection des activités des piétons [Enzweiler and Gavrilu, 2009] ou encore proposant une classification selon la reconnaissance d'activités d'un groupe de personnes [Aggarwal and Ryoo, 2011]. Comme illustré par la figure 6.1, nous nous intéressons plus spécifiquement à la reconnaissance d'activités simples, c'est à dire ne faisant pas intervenir une séquence longue d'événements, que nous cherchons à reconnaître à l'aide d'une caméra fixe, positionnée dans un environnement relativement peu maîtrisé (l'intérieur d'un habitat, avec toute la diversité des aménagements possibles). Nous reprenons, pour positionner notre contribution, la classification réalisée par Turaga *et al.* [Turaga *et al.*, 2008] qui regroupent les méthodes couramment utilisées pour la détection d'activités simples en trois catégories : les approches non paramétriques, volumétriques et paramétriques.

6.1.1 Approches non paramétriques

Ces approches consistent à récupérer une série de caractéristiques à chaque image de la vidéo. Cela permet d'obtenir une séquence de caractéristiques qui sera alors comparée aux séquences de caractéristiques des exemples types qui ont été pré-enregistrés. L'exemple pré-enregistré le plus similaire est considéré comme représentatif de l'activité réalisée par la personne. Cette approche peut être utilisée pour des objets extraits en 2 ou 3 dimensions.

6.1.2 Approches volumétriques

Ces approches consistent à analyser en un bloc l'enregistrement vidéo de l'activité réalisée par une personne. L'idée est d'analyser la séquence d'images en une seule fois et non pas image par image. A partir d'une séquence d'images, un volume en trois dimensions est créé à partir de l'ensemble des pixels appartenant à la personne. Chaque image constitue une coupe de ce

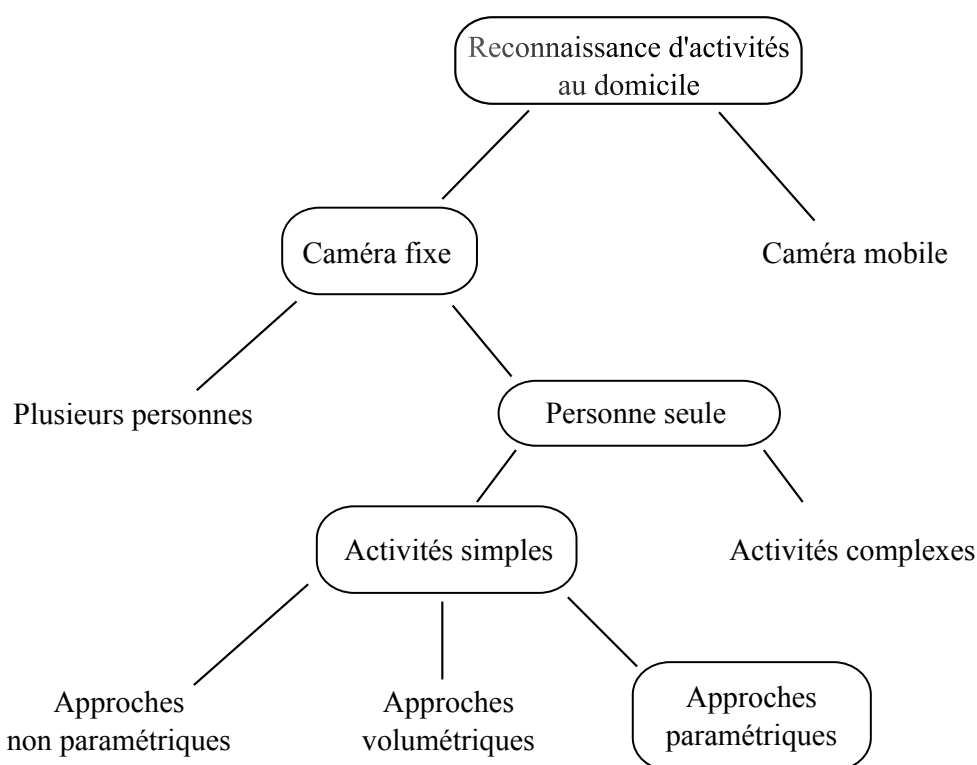


FIGURE 6.1 – Positionnement du problème de la reconnaissance d'activité au domicile.

volume qui représente le trajet effectué par la personne durant la séquence. À partir de ce volume, des caractéristiques sont extraites. Ces caractéristiques peuvent ensuite être comparées aux caractéristiques des exemples pré-enregistrés.

6.1.3 Approches paramétriques

Les approches paramétriques consistent à construire un modèle à partir d'hypothèses faites sur la dynamique de l'activité à reconnaître. Les paramètres spécifiques d'une activité sont appris à partir d'une base d'exemple. Les séquences d'images à reconnaître sont ensuite confrontées aux modèles des différentes activités. Le modèle le plus représentatif de la séquence permet d'identifier l'activité réalisée.

6.2 Positionnement

L'objectif est de détecter les chutes et situations à risque mais aussi de reconnaître des activités ou postures que peut prendre la personne dans la vie de tous les jours, en différenciant les activités où la personne est passive et celles où la personne est active. Les activités ont été choisies en s'inspirant des travaux de Noury *et al.* [Noury *et al.*, 2007]. Dans cet article, les auteurs indiquent comment construire un système détectant les chutes. Notamment, ils précisent que pour détecter les chutes, il est nécessaire de confronter le système à des situations réalistes pouvant se confondre aux situations de chute, qui peuvent entraîner des faux positifs. Ils précisent qu'il ne suffit pas de comparer la chute à des activités telles que la marche, mais de prendre également des situations telles que se pencher ou s'accroupir. Ils énumèrent ainsi une liste d'activités à prendre

en compte pour détecter efficacement les chutes. Nous avons identifié les activités

- à risque : chuter et être allongé au sol, monter sur un obstacle ;
- passives : être assis, allongé sur un lit ou sur un canapé ;
- actives : marcher, s'accroupir, se pencher, debout.

Les approches non paramétriques et volumétriques requièrent un grand nombre d'exemples pour rendre la détection d'activité la plus fiable possible. Dans notre cas, nous sommes partis du principe que notre base de données ne peut pas être exhaustive. En effet, il est difficile d'enregistrer toutes les activités ou les postures qu'une personne peut réaliser au cours de la journée. Enregistrer les mouvements naturels et complètement réalistes d'une personne est très difficile à faire en situation de laboratoire. Nous ne disposons, par exemple, pas d'enregistrements de chutes réelles. De plus, une même activité peut se différencier par l'angle de vue de la caméra, l'endroit où elle est réalisée ou la façon qu'ont les personnes de la réaliser. Les personnes âgées peuvent se comporter différemment du fait de leurs capacités motrices plus ou moins conservées. Nous avons donc utilisé des modèles de Markov cachés qui font partie des approches paramétriques et qui ont de bonnes capacités de généralisation à partir d'une base de données restreinte. En plus de la question de la reconnaissance de l'activité, ces modèles permettent de traiter le problème de la segmentation des activités dans les séquences d'images, alors que les autres approches nécessitent de renseigner le début et la fin de l'activité.

Dans la littérature, les articles traitant la détection d'activité à l'aide de méthodes paramétriques utilisent généralement un grand nombre de paramètres extraits des images. Par exemple, dans l'article de Peursum *et al.* [Peursum *et al.*, 2005], les auteurs décrivent leur méthode pour reconnaître l'action « imprimer » ou « préparer un thé » résultant d'une succession de 15 activités comme « marcher », « s'asseoir », etc. Pour reconnaître ces activités, ils se sont basés sur 10 caractéristiques à extraire de chaque image, comme la taille de la personne, la vitesse horizontale, la longueur du torse, l'angle du torse par rapport au sol, la longueur des mains, la longueur des jambes, etc. L'extraction de ces caractéristiques nécessite d'effectuer un certain nombre de traitements comme extraire la personne, localiser sa main, son torse, ses jambes et mesurer les angles au niveau des articulations.

Comme pour le choix des états, les caractéristiques ont été choisies en se basant sur l'article de Noury *et al.* [Noury *et al.*, 2007]. Les auteurs ont réalisé un état de l'art sur les méthodes utilisées pour développer un système de détection de chute. Ainsi, 4 approches sont couramment utilisées (séparément) pour détecter une chute d'une autre activité :

- détecter la position allongée sur le sol ;
- détecter l'impact du choc lorsque la personne se cogne au sol ou dans un obstacle, représenté par l'inversion de polarité du vecteur d'accélération dans la direction de la trajectoire ;
- détecter un manque de mouvement ;
- détecter la vitesse verticale, augmentant linéairement, due à l'accélération gravitationnelle lors de la chute.

Chaque approche prise séparément peut provoquer des fausses détections. Notre idée est de coupler les sources d'informations pour établir une décision plus robuste et reconnaître des activités pouvant se ressembler (comme se pencher et s'accroupir). Nous avons repris l'idée qu'une chute se produit, lorsque le corps de la personne a une vitesse verticale qui augmente et que la personne se trouve allongée au sol. Pour obtenir ces informations, nous prenons en compte, dans un premier temps, trois caractéristiques, la vitesse verticale du centre de masse, la position du centre de masse et la dispersion verticale du corps de la personne (l'écart-type) sur l'axe vertical, auxquelles nous ajouterons ensuite la vitesse horizontale et la position de la tête. La vitesse verticale du centre de masse informe sur le déplacement du corps à chaque instant. La position

du centre de masse permet de savoir si la personne est proche du sol. La distribution verticale du corps renseigne sur la forme du corps, ainsi il sera possible de distinguer des activités, comme s'allonger et s'accroupir, qui sont toutes deux des activités où la position du centre de masse est proche du sol et où la vitesse verticale est nulle, mais qui auront des dispersions verticales différentes.

Nous proposons une approche nécessitant peu de caractéristiques pour reconnaître l'activité d'une personne tout en étant capable de généraliser à partir d'une base de données restreinte. En effet, la base de données sur laquelle nous travaillons est formée à partir de différentes situations réalisées par 28 sujets jeunes. Ces situations correspondent aux activités que nous voulons détecter. Plus d'informations sur la constitution de cette base est donnée au chapitre 8.1. Il s'agit de trouver un compromis minimisant la complexité du modèle et en particulier le nombre d'observations tout en garantissant une bonne robustesse dans la reconnaissance. Nous proposons plusieurs modèles dont les paramètres peuvent être ajustés par apprentissage mais avec relativement peu de données. Certains de ces modèles permettent de prendre en compte une connaissance de bon sens par la définition manuelle d'une partie des paramètres.

Dans la suite de ce chapitre, nous présentons différents modèles d'activité que nous avons évalués. Ils diffèrent par le nombre d'activités modélisées, le nombre d'observations et par l'insertion ou non de connaissance *a priori*. Ces différents modèles sont proposés dans le but d'être ensuite évalués expérimentalement afin d'identifier le niveau de complexité permettant d'obtenir une reconnaissance précise des activités tout en étant suffisamment général pour permettre la reconnaissance de situations éloignées de la base d'apprentissage.

6.3 Proposition d'un modèle de l'activité

Les modèles de Markov cachés (MMC) fournissent un cadre efficace pour modéliser des phénomènes ou processus régis par une dynamique d'états cachés. Seules des observations partielles sur le processus sont accessibles et les états peuvent être inférés avec une certaine probabilité à partir d'une séquence d'observations. La figure 6.2 illustre un MMC évoluant dans le temps avec X_t les états cachés, représentés en gris, et O_t les observations. Un MMC est défini formellement par :

- un ensemble d'états ;
- une matrice $a_{ij} = P(X_t = j | X_{t-1} = i)$ représentant toutes les probabilités de transition entre chaque paire d'états ;
- une probabilité initiale $\pi_i = P(X_0 = i)$ pour chaque état i , qui est la probabilité *a priori* que le système se trouve dans l'état i à $t = 0$;
- une fonction d'observation $b_i(\mathbf{o}) = P(O_t = \mathbf{o} | X_t = i)$.

Des algorithmes d'apprentissage et d'inférence bien connus sont présentés en annexe A. L'objectif est de construire un modèle permettant de reconnaître l'activité réalisée par la personne filmée.

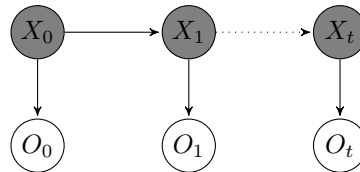


FIGURE 6.2 – Représentation graphique d'une chaîne de Markov cachée, avec X_t le processus Markovien caché et O_t la séquence d'observations correspondantes.

Le modèle le plus simple que nous proposons est constitué d'un état par activité avec pour chaque état une distribution gaussienne formant la fonction d'observation. Les paramètres de ces distributions sont appris à partir de la base de données. Nous ajoutons ensuite à ce modèle de référence une fonction de transition définie manuellement qui nous permet d'injecter de la connaissance sur l'enchaînement des activités. La fonction de transition est définie manuellement car nous ne disposons pas de données représentatives des enchaînements possibles des situations que nous voulons reconnaître. Cette seconde proposition prend donc la forme d'un MMC unique permettant de reconnaître les différentes activités avec un état par activité. Enfin nous proposons d'utiliser plusieurs MMC, un MMC pour chaque activité dont les paramètres sont appris à partir de la base de données. Dans cette dernière proposition les transitions entre activités ne sont pas définies. Il n'y a donc pas de contraintes *a priori* sur les enchaînements ce qui nous permet d'avoir un modèle plus général.

6.3.1 Modèle avec fonction d'observation seule

Etats Le modèle contient 8 états correspondant à 8 activités qui sont : s'asseoir, marcher, s'accroupir, se pencher, monter sur un obstacle (une chaise, une marche, etc.), s'allonger sur un lit ou sur un canapé, chuter et être allongé au sol. Dans ce modèle, les activités « marcher » et « être debout » sont confondues.

Observations Nous utilisons les caractéristiques suivantes :

- la position du centre de masse sur l'axe vertical ;
- la vitesse verticale du centre de masse ;
- la dispersion verticale du corps de la personne (l'écart-type) sur l'axe vertical.

La fonction d'observation, donnant la probabilité d'observer une valeur $\mathbf{o}$ dans l'état i , suit une loi normale multidimensionnelle pour chaque valeur de i :

$$P(O = \mathbf{o} | X = i) = \frac{1}{(2\pi)^{N/2} |\Sigma_i|^{1/2}} e^{-1/2(\mathbf{o} - \mu_i)^T \Sigma_i^{-1} (\mathbf{o} - \mu_i)} \quad (6.1)$$

avec Σ_i la matrice de covariance, μ_i la moyenne et N le nombre de dimensions (pour ce MMC le N est égal à 3, car il y a trois observations). Σ_i et μ_i sont calculés, pour chaque état i , par apprentissage à partir de la base de données. Après apprentissage, nous obtenons donc 8 lois normales, une pour chaque état.

Conclusion Ce modèle sans transition est représenté à la figure 6.3. Sur cette figure, les huit carrés représentent les huit états de notre modèle. Les 8 états représentent les 8 activités réalisées par le sujet, détecté en vert par l'algorithme. Cette figure montre qu'aucune transition n'est définie entre les états. La décision de l'état dans lequel se trouve la personne repose, ici, essentiellement sur les observations.

Inférence

Pour réaliser l'inférence, l'état X_t le plus probable est donnée à chaque pas de temps en comparant les valeurs prises, pour l'observation $\mathbf{o}_t$, par les 8 distributions gaussiennes. Le calcul est le suivant :

$$P(X_t = i | O = \mathbf{o}_t) = \frac{P(O = \mathbf{o}_t | X_t = i) P(X_t = i)}{\sum_i P(O = \mathbf{o}_t | X_t = i) P(X_t = i)}.$$

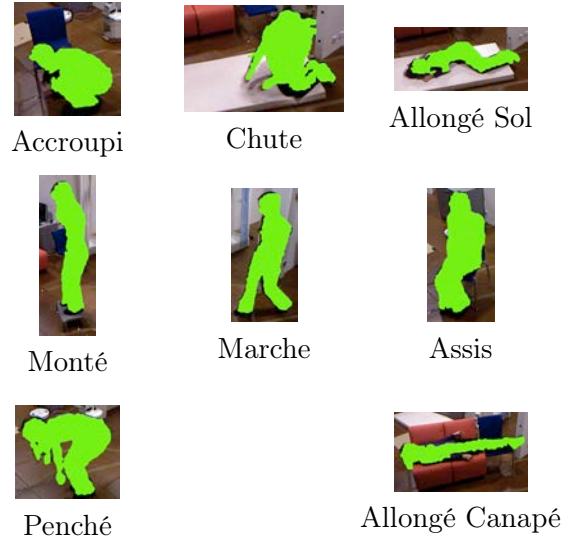


FIGURE 6.3 – Modèle à 8 états avec fonction d'observation seule.

Nous n'avons pas considéré de différence *a priori* entre les états. La probabilité $P(X_t = i)$ est donc uniforme pour toutes les valeurs de i et l'équation précédente se résume donc à :

$$P(X_t = i | O = \mathbf{o}_t) = \frac{P(O = \mathbf{o}_t | X_t = i)}{\sum_i P(O = \mathbf{o}_t | X_t = i)}.$$

6.3.2 Utilisation d'un MMC unique

Nous ajoutons à la proposition précédente, qui associe un état à chaque activité, une matrice de transition qui modélise l'enchaînement dans le temps des activités. Nous obtenons donc un MMC représentant l'activité comme un processus dynamique partiellement observable. Nous proposons dans cette section un premier modèle à 8 états et un second à 9 états. Pour chaque modèle, nous définissons l'espace d'états (les activités à reconnaître), la probabilité initiale qu'à le processus (l'activité) d'être dans un état donné, la dynamique (transition entre états) ainsi que les observations considérées. Dans ces modèles, la fonction d'observation est apprise à partir de la base de données. La matrice de transition est, quant à elle, définie manuellement. Nous partons en effet du principe qu'il n'est pas possible d'apprendre les transitions entre activités à partir d'une base de donnée constituée en laboratoire. Cette base de donnée, constituée à partir de situations jouées, ne peut pas être représentative de la diversité des enchaînements d'activité d'une personne dans la vie de tous les jours. Il est en revanche possible d'injecter une connaissance *a priori*, de bon sens, par la définition manuelle des probabilités de transition.

MMC à 8 états et 3 observations

Etats Les états sont les mêmes que précédemment, à savoir : s'asseoir, marcher, s'accroupir, se pencher, monter sur un obstacle (une chaise, une marche, etc.), s'allonger sur un lit ou sur un canapé, chuter et être allongé au sol.

Probabilités initiales Nous considérons qu'au début de l'analyse la personne peut se trouver dans n'importe quel état. Nous attribuons donc la même probabilité initiale pour chaque état,

soit 0,125.

Matrice de transition L'activité est modélisée comme un processus stochastique. Les transitions entre les différents états expriment la dynamique de l'activité qui tient compte autant des contraintes biomécaniques que d'un *a priori* sur le comportement des personnes. Pour insérer les transitions possibles, nous avons fait plusieurs hypothèses.

- La journée d'une personne âgée, à son domicile, peut être résumée en une succession de passages assis-debout. Nous définissons les transitions les plus probables comme étant celles qui mènent vers l'état « Marche ».
- La personne marche (ou est debout) juste avant et après avoir effectué les activités « Accroupi », « Monté », « Penché », « Assis ».
- Une personne ne peut pas directement passer de la position debout à allonger mais elle doit d'abord s'asseoir pour pouvoir s'allonger.
- La probabilité d'aller dans l'état « Chute » est très faible comparée aux autres activités.
- Une personne se trouvant allongée au sol est obligatoirement passée par l'état « Chute ».
- Une fois au sol, la personne peut soit y rester soit, dans le cas où la personne est capable de se relever, passer directement à l'état « Marche » ou passer par l'état « Accroupi » qui lui permettra de passer par l'état « Marche » (ou être debout) par la suite.

Le MMC construit sur la base de ces hypothèses est représenté à la figure 6.4. Les transitions correspondent aux flèches et les états aux rectangles. Nous définissons les probabilités relatives des différentes transitions sur une échelle à quatre niveaux :

- pp , la plus petite probabilité de transition ;
- p , une probabilité de transition faible ;
- m , une probabilité de transition moyenne ;
- C , la probabilité majoritaire correspondant au complément à 1 des autres probabilités sortantes de l'état.

Observations Le modèle construit ici reprend les 3 mêmes observations que le précédent modèle, à savoir la position et la vitesse verticale du centre de masse et la distribution du corps sur l'axe vertical. La fonction d'observation est également conservée, elle suit une loi normale multidimensionnelle.

Conclusion Avec la construction de ce MMC, nous injectons de la connaissance à travers la matrice de transition.

MMC à 9 états et 5 observations

Dans l'objectif de mesurer les paramètres de la marche d'une personne à son domicile et donc de pouvoir appliquer au domicile les résultats présentés au chapitre 5, il est important d'isoler précisément les moments où la personne se déplace. Cela implique de différencier l'activité « Marche » de l'activité « Debout ».

Etats Dans ce nouveau modèle, nous ajoutons un état « Debout » aux huit états précédents.

Probabilités initiales Comme pour les précédents modèles, nous attribuons une même probabilité de $1/9$ à chaque état.

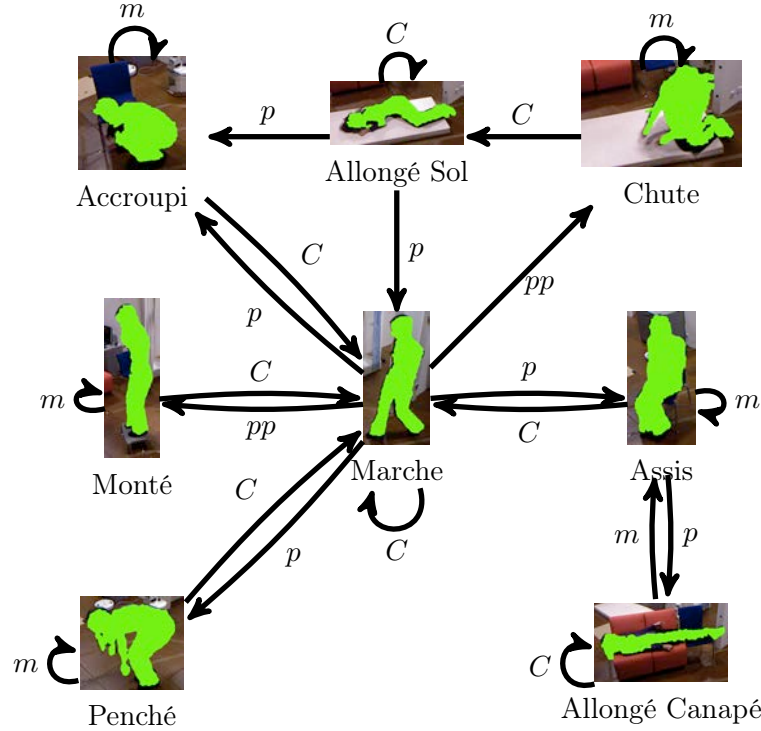


FIGURE 6.4 – MMC à 8 états avec des transitions définies manuellement.

Matrice de transition Les transitions et les probabilités sont définies sur la base des mêmes hypothèses que celles du modèle précédent, en ajoutant :

- l'état « Debout » précède et suit la plupart des activités. Seules les activités « Monté » et « Chute » peuvent être aussi précédées et suivies de l'état « Marche ».

Le MMC construit sur la base de ces hypothèses est représenté à la figure 6.4. Les probabilités relatives des différentes transitions sont définies sur la même échelle à quatre niveaux que précédemment.

Observation L'ajout de l'état « Debout » nécessite une augmentation du nombre d'observations. En effet, les trois observations précédentes ne suffisent pas à différencier les états « Debout » et « Marche ». Intuitivement, nous pouvons penser que la vitesse horizontale diffère entre les deux états. Nous ajoutons également la position verticale de la tête pour améliorer la robustesse du modèle. Ainsi, dans ce modèle nous insérons 5 observations : la position du centre de masse, la dispersion verticale du corps, la vitesse verticale et la vitesse horizontale du centre de masse et le point maximum du corps. La fonction d'observation suit toujours une loi normale multidimensionnelle avec cette fois un nombre de dimensions égal à 5 soit $N = 5$ dans la formule 6.1. Comme précédemment, à partir de la base de données, les moyennes μ des cinq observations et la covariance Σ sont calculées pour chacun des neuf états. Nous obtenons ici 9 lois normales.

Conclusion Nous voulons ici démontrer que nous pouvons affiner les activités réalisées, notamment distinguer la personne active (marchant) de la personne statique. La contrepartie est d'augmenter la complexité du modèle en augmentant le nombre des observations en passant de 3 à 5.

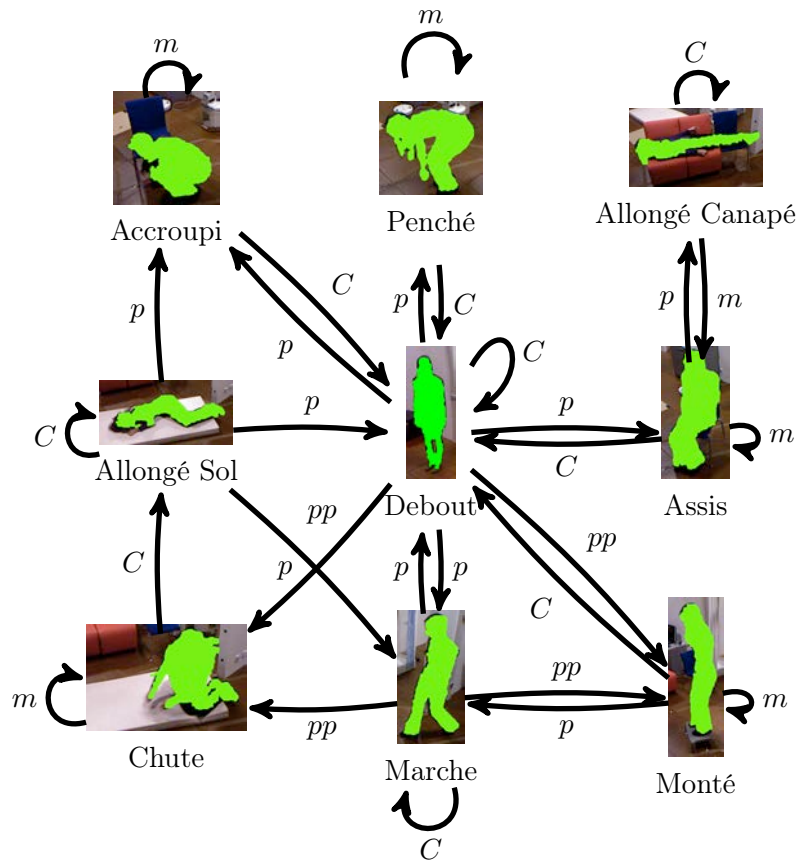


FIGURE 6.5 – MMC à 9 états avec des transitions définies manuellement.

Inférence

Pour savoir dans quel état se trouve la personne, nous avons implémenté les algorithmes *Forward-Backward* et de Viterbi, décrit en annexe A. En implémentant les deux algorithmes, les plus couramment utilisés avec les MMC, nous avons voulu savoir si des différences notables pouvaient apparaître dans la décision finale de l'état dans lequel se trouve la personne. La comparaison des deux algorithmes est présentée dans le chapitre 8 lors de la présentation des résultats de l'expérimentation.

6.3.3 Utilisation de MMC combinés

Dans les modèles précédents, chaque activité correspondait à un état et nous obtenions donc un MMC à 8 ou 9 états selon le modèle. Avec l'approche multi-modèles, chaque activité correspond à un MMC, nous obtenons donc 8 ou 9 MMC selon le nombre d'activités à reconnaître. Une autre différence avec les modèles précédents est que nous n'insérons aucune connaissance *a priori* (ni transition, ni probabilité initiale). Comme dans le cas du modèle avec les observations seules, la dynamique entre les situations n'est pas prise en compte.

Système de 8 MMC

Etat Nous construisons deux types de systèmes de modèles. Ces deux systèmes de modèles sont identiques à la seule différence que l'un est constitué de 8 MMC à 2 états et l'autre à trois états chacun. Nous voulions ainsi tester l'influence qu'a le nombre d'état choisi.

Probabilités initiales et matrice de transition Comme nous venons de le préciser, aucune connaissance n'est insérée dans ce modèle. Cela signifie que nous ne définissons pas manuellement les probabilités de transition et les probabilités initiales. Ces probabilités sont apprises automatiquement à partir de la base de données.

Observations Les observations sont les mêmes que pour les MMC à 8 états, à savoir la position et la vitesse du centre de masse sur l'axe vertical et la distribution du corps sur l'axe vertical. Chacun des huit MMC se base sur ces trois observations. La fonction d'observation reste elle aussi inchangée, elle suit une loi normale multidimensionnelle. Les paramètres de cette loi normale sont calculés à partir de la base de données. Les paramètres sont appris pour chaque état des 8 MMC. Nous obtenons donc 16 ou 24 lois normales selon le modèle à deux ou trois états par MMC.

Conclusion Le système des 8 MMC à deux états chacun et sans probabilité de transition est représenté à la figure 6.6. Nous appliquons à ces systèmes de modèle l'algorithme de Baum-Welch permettant d'apprendre par lui-même les paramètres du MMC, à partir de la base de données. Les paramètres qui sont appris sont les probabilités de transition entre chaque état, les probabilités initiales et les lois normales pour chaque état. Le nombre d'états pour chacun des 8 MMC est le seul paramètre que nous déterminons à la main.

Système de 9 MMC

Etat Comme précédemment, nous construisons deux types de systèmes de modèles. C'est-à-dire que nous développons un système de 9 modèles à 2 états et un autre système à trois états chacun.

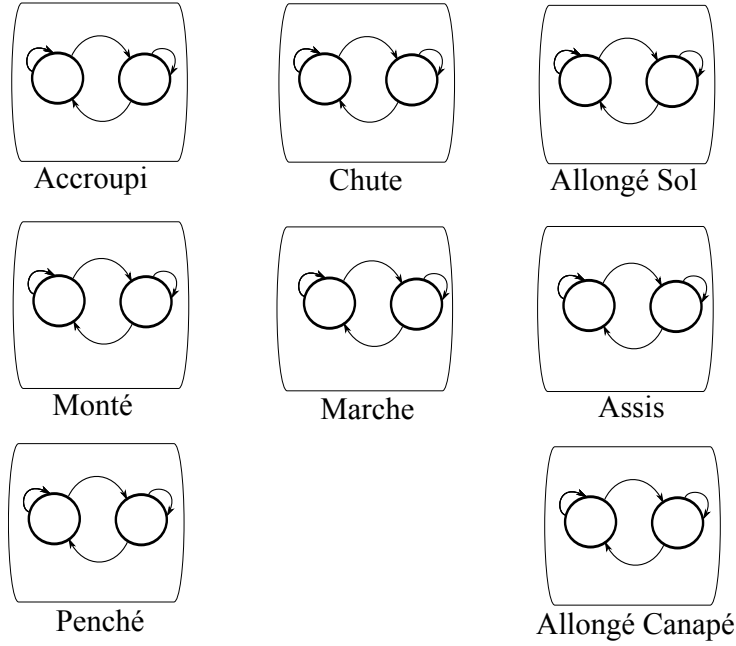


FIGURE 6.6 – Système de 8 MMC à deux états chacun.

Probabilités initiales et matrice de transition Comme précédemment aucune connaissance n'est insérée dans ce modèle. Cela signifie que les probabilités initiales et les matrices de transition sont apprises automatiquement à partir de la base de données.

Observations Les observations sont les mêmes que pour les MMC à 9 états, à savoir la position et la vitesse du centre de masse sur l'axe vertical, la dispersion du corps sur l'axe vertical, la vitesse horizontale du centre de masse et la position du point maximum du corps. Chacun des neuf MMC se base sur ces cinq observations. La fonction d'observation reste elle aussi inchangée, elle suit une loi normale multidimensionnelle. Les paramètres de cette loi normale sont calculés à partir de la base de données. Les paramètres sont appris pour chaque état des 9 MMC. Nous obtenons donc 18 ou 27 lois normales selon le modèle à deux ou trois états par MMC.

Conclusion Le système des 9 MMC à deux états chacun et sans probabilité de transition est représenté à la figure 6.7. Nous appliquons à ces systèmes de modèle, comme pour les 8 MMC, l'algorithme de Baum-Welch permettant d'apprendre par lui-même les paramètres du MMC à partir de la base de données. Les paramètres, qui sont appris, sont les probabilités de transition entre chaque état, les probabilités initiales et les lois normales pour chaque état. Le nombre d'états pour chacun des 9 MMC est donc le seul paramètre que nous déterminons à la main.

Inférence

Pour savoir dans quel état se trouve la personne, nous devons d'abord appliquer un algorithme d'apprentissage, avant d'utiliser les algorithmes d'inférence, comme décrit en annexe A. Nous avons donc utilisé l'algorithme de Baum-Welch pour apprendre les paramètres de chaque modèle à partir de la base de données que nous avons construite. L'inférence est réalisée, avec l'algorithme *Forward*, en calculant à chaque pas de temps la vraisemblance des différents modèles sur les

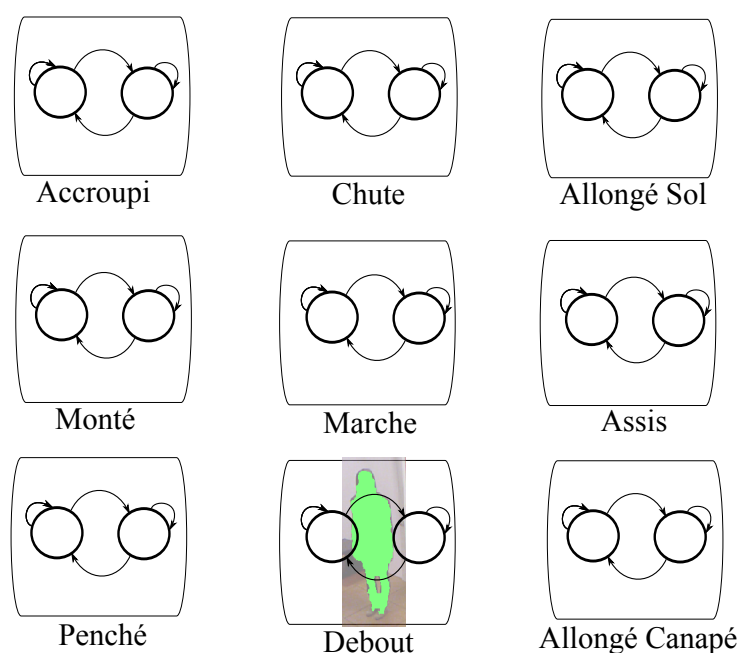


FIGURE 6.7 – Système de 9 MMC à deux états chacun.

dernières images de la séquence. L'activité la plus probable est celle qui correspond au modèle donnant la plus grande vraisemblance.

6.4 Conclusion

Dans ce chapitre, nous avons présenté les différents modèles que nous avons construits pour reconnaître l'activité. Chaque modèle a des avantages et inconvénients. Certains modèles sont plus simples en ne prenant que trois observations (les modèles à 8 états ou 8 MMC combinés) mais ne font, par conséquent, pas la différence entre être debout et marcher (comme c'est le cas avec les modèles à 9 états ou 9 MMC combinés). Certains modèles n'ont pas de connaissance *a priori* et reposent sur un nombre plus grand de paramètres à apprendre, ce qui implique de travailler avec une base de données plus importante comparée aux modèles où l'on insère les paramètres manuellement en utilisant une connaissance humaine de l'enchaînement des situations. Les résultats donnant les meilleurs taux de classification sont présentés dans le chapitre 8.

Conclusion

Nous avons identifié la caméra de profondeur comme capteur répondant à nos besoins pour développer un système de maintien à domicile des personnes âgées. Un certain nombre de traitements sur le flux vidéo sont à effectuer pour, ensuite, pouvoir extraire de l'information de plus haut niveau, comme détecter l'activité ou analyser la marche d'une personne.

Le premier chapitre nous a permis de présenter les différentes méthodes en vision pour détecter la personne et la suivre au cours du temps sur l'ensemble des images fournies par la caméra. L'état de l'art n'est pas exhaustif, mais nous avons identifié les méthodes utilisées pour détecter et suivre une seule personne en intérieur avec une caméra fixe.

Le deuxième chapitre présente les algorithmes de détection de la personne et de suivi que nous avons développés. Nous avons opté pour des méthodes légères, ne nécessitant pas de temps de calcul importants pour que ces méthodes restent compatibles avec les objectifs de traitement bas coût et de temps réel, que nous nous sommes fixés. A partir des images de profondeur, fournies par la caméra, nous obtenons le fond par apprentissage, puis la silhouette de la personne est extraite par soustraction avec l'image de profondeur courante. La silhouette est représentée par un ensemble de points, dont sont extraits 5 paramètres : le centre de masse, la vitesse verticale et la vitesse horizontale du centre de masse, la dispersion verticale et le point maximum de la silhouette.

L'algorithme d'analyse de la marche, développé dans le troisième chapitre, nous permettra d'évaluer le degré de fragilité de la personne en analysant l'évolution des paramètres de la marche spatio-temporels (longueurs, cadence, vitesse de marche). Cet algorithme est très simple, puisqu'il se base uniquement sur l'information du centre de masse et la détection des maxima locaux sur l'axe vertical. Ainsi, le système peut fournir, en temps réel, les paramètres de la marche de la personne suivie. Il sera alors possible d'observer des évolutions des paramètres de la marche au cours du temps, pour déterminer des éventuels risques de perte d'autonomie. Actuellement, les médecins observent certains de ces paramètres en consultation et évaluent l'évolution lors des visites suivantes. Le problème étant qu'entre chaque visite, il peut se passer quelques mois. Or, il serait possible d'effectuer une analyse quotidienne des paramètres de la marche en installant notre système au domicile des personnes. Ainsi, l'évaluation de la fragilité d'une personne pourrait être faite à partir d'un plus grand volume d'informations.

Le quatrième chapitre est consacré à la présentation des algorithmes utilisés pour reconnaître l'activité d'une personne. Le but de cet axe de recherche est de sécuriser et évaluer la personne. Ainsi pour sécuriser la personne, nous voulons développer un système détectant la chute et les situations à risque telles que monter sur une chaise et s'accroupir, qui sont des activités pouvant conduire à une chute, en particulier chez les personnes âgées. Puis, pour évaluer la personne, nous voulons développer un système capable de déterminer si la personne réalise des activités actives, autres que s'asseoir et s'allonger. Ainsi, nous avons construit un algorithme qui identifie les activités suivantes : s'asseoir, marcher, s'accroupir, se pencher, monter sur un obstacle (une chaise, une marche, etc.), s'allonger sur un lit ou sur un canapé, chuter et être allongé au sol.

Parmi les méthodes existantes pour reconnaître les activités, nous avons choisi de développer un algorithme appartenant à la catégorie des méthodes paramétriques. L'une des difficultés, en reconnaissance d'activité, est la variabilité des données, du fait de la différence entre les personnes, de la façon qu'elles ont de réaliser les activités, de l'angle de vue de la caméra, etc. Nous avons choisi les modèles de Markov cachés qui ont de bonnes capacités de généralisation. Dans ce quatrième chapitre, nous avons présenté différents modèles que nous avons construits pour reconnaître l'activité. Nous avons proposé deux approches. La première utilise un modèle unique dans lequel une connaissance *a priori* est injectée par la définition de la matrice de transition. Dans la seconde, plusieurs modèles sont construits par apprentissage et les transitions entre activités ne sont pas modélisées. Nous avons fait varier la complexité des modèles en jouant sur le nombre d'activité à reconnaître et sur le nombre d'observations. Nous avons construit des modèles pour reconnaître 8 activités, utilisant 3 observations, la position et la vitesse verticale du centre de masse et la distribution verticale de la silhouette. Nous avons construit d'autres modèles reconnaissant 9 activités en ajoutant l'activité « Debout ». Ces modèles discriminent les 9 activités à partir de 5 observations, les trois mêmes que précédemment, en ajoutant la vitesse horizontale du centre de masse et le point le plus haut du corps. Dans le cas des systèmes de MMC, nous avons implémenté l'algorithme de Baum-Welch pour apprendre les paramètres des modèles, puis nous avons implémenté *Forward* pour inférer l'état dans lequel se trouve la personne. Dans le cas des MMC uniques, nous avons implémenté l'algorithme *Forward-Backward* et de Viterbi pour inférer l'état à partir de la séquence d'observation. Le point commun entre tous ces modèles est qu'ils sont fondés sur peu d'observations et que ces observations sont simples à extraire des images. Les différents modèles sont ensuite testés et comparés sur la base du nombre de bonnes classifications qu'ils effectuent. Les résultats sont présentés au chapitre 8.

Troisième partie

Evaluation du système

Introduction

Dans les deux chapitres précédents, nous avons présenté nos contributions à la reconnaissance d'activités et à l'analyse en milieu écologique des paramètres de la marche. Dans cette partie, nous en évaluons les performances. Plusieurs expérimentations sont mises en place. Les deux premiers chapitres présentent des expérimentations qui ont été menées dans un appartement expérimental avec des sujets jeunes. Le dernier chapitre présente quelques résultats obtenus avec des personnes âgées.

Le premier chapitre décrit une expérimentation qui a été menée pour vérifier les performances du dispositif d'analyse de la marche que nous proposons, en comparaison avec les mesures fournies par un tapis actimétrique que nous considérons comme la vérité terrain. Dans cette expérimentation, nous testons la précision pour différentes situations, pouvant être rencontrées dans la réalité. Par exemple, nous avons demandé aux personnes de marcher perpendiculairement à la caméra, puis de face, ensuite nous leur avons demandé de faire des petits pas comme pourraient le faire des personnes âgées. Les résultats nous montrent que le système est précis lorsque la personne marche perpendiculairement à la caméra, et ce pour toutes les situations.

Le deuxième chapitre teste l'algorithme de détection des activités. Nous avons constitué une base de données à partir de laquelle, une partie des séquences sont conservées pour que l'algorithme apprenne, automatiquement, les paramètres de chaque MMC. Nous évaluons, ensuite, les modèles sur la base du nombre de bonnes classifications des séquences restantes. Les modèles à 8 états ou 8 MMC donnent les meilleurs résultats sur des séquences proches de celles constituant la base d'apprentissage, avec la personne entièrement visible. Nous testons les capacités de généralisation de l'algorithme avec une deuxième expérimentation. Nous présentons à l'algorithme des séquences où la personne n'est pas entièrement visible. Les résultats montrent que certains modèles sont robustes sur ces nouvelles séquences, ce qui est le cas pour le modèle à 9 états.

Le dernier chapitre de cette partie présente des tests qui ont été effectués dans un hôpital et au domicile, avec des personnes âgées. 6 personnes âgées ont marché devant notre système. Il ne s'agit pas d'une réelle expérimentation mais de premiers tests pour lesquels nous n'avons effectué qu'une analyse qualitative des résultats. Concernant l'algorithme de la marche, les résultats montrent que, malgré l'irrégularité des trajectoires du centre de masse, comparées aux trajectoires que nous obtenions avec des sujets plus jeunes, l'algorithme réussit à extraire des paramètres de la marche. Concernant la détection des activités, seules quelques activités ont été testées. Parmi celles-ci, la marche, la position assise et la position penchée ont été reconnues.

Chapitre 7

Expérimentations pour l'analyse de la marche

Sommaire

7.1	Première évaluation du système	89
7.1.1	Description de l'expérimentation	89
7.1.2	Résultats	91
7.1.3	Conclusion	94
7.2	Évaluation globale du système	94
7.2.1	Description de l'expérimentation	94
7.2.2	Résultats	95
7.3	Discussion et conclusion	97

L'expérimentation a pour but de tester la précision des paramètres de la marche fournis par notre algorithme, lorsqu'une personne marche dans le champ de vision de la caméra. Une première évaluation a tout d'abord été réalisée. L'objectif est de vérifier que l'algorithme est assez robuste pour commencer l'évaluation et de valider ou non certaines hypothèses. A la suite des résultats obtenus avec cette première évaluation, le protocole expérimental d'une seconde expérimentation a été construit.

7.1 Première évaluation du système

Cette section présente la première évaluation basée sur le papier et l'encre et qui constitue notre référence pour évaluer la précision de l'algorithme à extraire les indicateurs de la marche. Nous présentons le protocole expérimental et les résultats.

7.1.1 Description de l'expérimentation

5 sujets sains ont intégré l'expérience dont une femme et quatre hommes, âgés de 20 à 53 ans. L'expérience consiste à placer, sur environ $4 m^2$, du papier blanc au sol (figure 7.1(b)). Nous avons placé des tampons imbibés d'encre sur la pointe et le talon des chaussures du sujet, comme montré à la figure 7.1(a). Ainsi lorsque le sujet marche sur le papier, il le marque de ses pas. Des méthodes similaires de marquage sur le sol peuvent être trouvées dans la littérature. Notamment, Georges Gilles de la Tourette a utilisé une technique similaire pour enregistrer la marche [André *et al.*, 2001]. Cette méthode a l'avantage d'être simple à mettre en place, peu

coûteuse et de ne nécessiter aucun équipement spécifique. Dans le cadre de l'expérience, la méthode « papier/encre » est considérée comme l'étalon permettant d'obtenir la vérité terrain. L'unique variable prise en compte dans cette étude est la longueur de pas. Pour obtenir les longueurs de pas, l'expérimentateur mesure manuellement la distance entre la marque du talon droit et celle du talon gauche sur le papier. Ces longueurs de pas sont alors comparées aux longueurs fournies par notre approche fondée sur une caméra RGB-D.



(a) Tampons imbibés d'encre sous les chaussures. (b) Un sujet marche sur le papier et le marque.

FIGURE 7.1 – Méthode utilisée comme référence pour mesurer les longueurs de pas.

Le sujet marche perpendiculairement à la caméra. Sachant que les personnes âgées font des plus petits pas [Hageman and Blanke, 1986] que les sujets testés dans notre étude, nous évaluons la précision de l'algorithme en le sollicitant sur un large échantillon de longueurs de pas. Les sujets réalisent les deux situations suivantes :

- marcher en ligne droite « normalement » ;
- marcher en ligne droite en faisant des petits pas.

Cette première évaluation a plusieurs objectifs :

1. évaluer si les résultats sont invariants à l'angle de la caméra,
2. évaluer la précision des mesures en fonction des différents algorithmes de filtrage choisis.

Nous voulons savoir si l'angle de la caméra a une incidence sur les résultats. C'est pourquoi les sujets réalisent les deux situations citées précédemment deux fois, une fois avec un angle égal à -27° et l'autre à 0° . Le placement de la caméra avec un angle à -27° correspond à la caméra positionnée en haut de la figure 7.2 et la caméra avec un angle à 0° correspond à la caméra positionnée en bas de la figure 7.2. Lors de l'expérience, une seule caméra enregistre à la fois. Nous voulons aussi tester la précision des longueurs de pas fournies par l'algorithme en fonction du filtre de lissage de la trajectoire du centre de masse sélectionné. A la section 4.2.2, deux méthodes de lissage des trajectoires ont été présentées, une méthode utilisant le filtre de Kalman et une autre utilisant un filtre passe-bas. Nous voulons savoir si la précision de l'extraction des maxima locaux, permettant d'obtenir les longueurs de pas, dépend du filtre utilisé. Pour tester la précision de notre algorithme, nous comparons les longueurs de pas fournies par le papier avec les longueurs fournies par l'algorithme utilisant soit le filtre de Kalman, soit le filtre passe-bas.

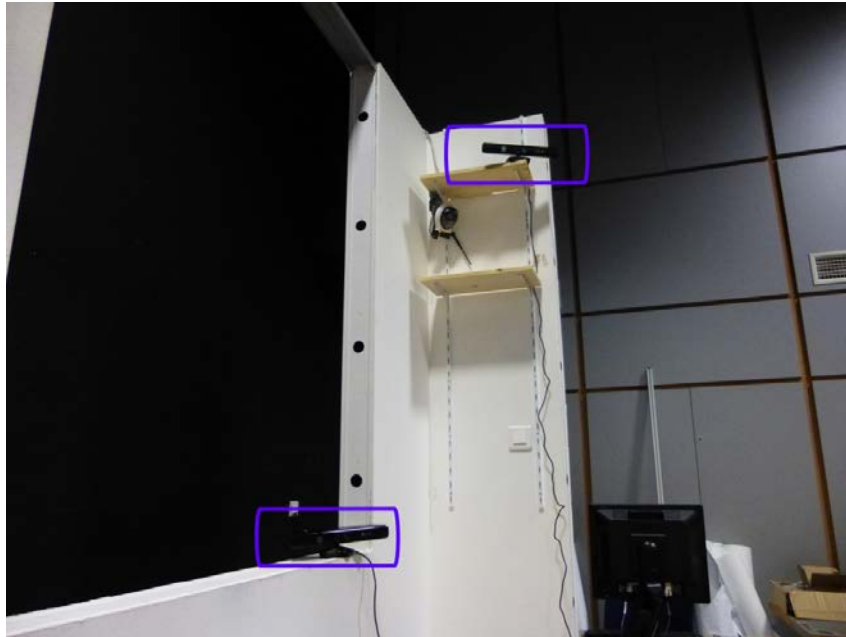


FIGURE 7.2 – Positionnement des caméras : la plus basse avec un angle de 0° et la plus haute avec un angle de -27° .

7.1.2 Résultats

Pour réaliser les différents tests statistiques et les graphes, nous avons utilisé le logiciel R (qui est un langage et un environnement pour les calculs statistiques).

Angle de la caméra

Afin de savoir si l'angle de la caméra modifie la qualité des résultats, les erreurs moyennes des longueurs de pas, obtenues quand la caméra a un angle à -27° et à 0° , sont calculées. Plus précisément, nous calculons l'erreur entre les longueurs fournies par le papier et par notre algorithme selon l'angle de la caméra, pour chaque pas. L'erreur moyenne est calculée comme la moyenne sur toutes les erreurs, considérant d'un côté les données pour la caméra à -27° et de l'autre les données pour la caméra à 0° . Les résultats sont montrés à la Table 7.1. L'erreur moyenne est donnée en centimètres et en pourcentages. Les erreurs moyennes sont similaires entre la caméra à -27° et à 0° . Comme indiqué par la table, l'erreur moyenne avec l'angle à -27° est de 5,4 cm soit une erreur moyenne de 11,6 %. Ce résultat est proche de l'erreur moyenne de 5,2 cm, soit une erreur moyenne de 11,9 %, obtenue avec l'angle à 0° . Un test statistique complète ces résultats, dans le but de savoir si les erreurs sur les longueurs de pas obtenues avec les deux angles sont significativement différentes. Avec un test de Student, pour deux échantillons indépendants, la p-value obtenue est de 0,4777. La p-value étant plus grande que 0,05, nous ne pouvons pas en conclure qu'il existe une différence significative entre les résultats fournis par la caméra positionnée à -27° et celle positionnée à 0° . Pour finir, le graphe, de la figure 7.3, montre que la variabilité des erreurs sur les longueurs de pas est la même pour l'angle à 0° que pour l'angle à -27° . Cela est montré par la forme des deux boîtes à moustache. La conclusion, que nous pouvons faire, est que notre algorithme est robuste à l'angle de la caméra choisi.

	Angle -27 °	Angle 0 °
Erreur moyenne (cm)	5,4	5,2
Erreur moyenne (%)	11,6	11,9
Nombre de pas	56	50

TABLE 7.1 – Comparaison des résultats, entre le tapis et la caméra, en fonction de l'angle de la caméra.

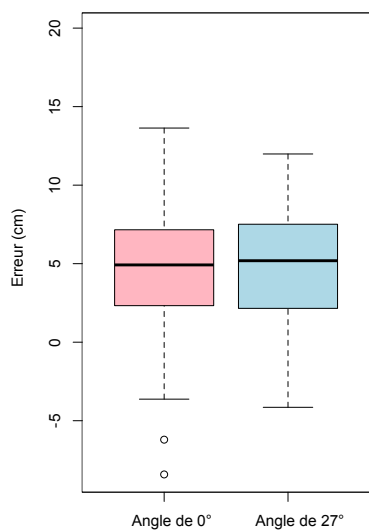


FIGURE 7.3 – Représentation des erreurs sur les longueurs de pas obtenues à partir d'une caméra avec un angle de -27 ° et avec un angle de 0 °.

Choix du filtre

Dans cette partie, le système est évalué sur la précision des longueurs de pas qu'il fournit, en fonction du filtre de lissage de la trajectoire du centre de masse. La trajectoire qui est correctement filtrée, nous permet de localiser plus précisément les maxima locaux, et ainsi nous permet d'en déduire les longueurs de pas les plus justes. Nous voulons ainsi tester la précision des longueurs de pas lorsque l'algorithme utilise le filtre de Kalman et le filtre passe-bas.

Suite à l'expérience effectuée sur 5 sujets, nous avons extrait les pas et selon le filtre utilisé, le filtre de Kalman ou le filtre passe-bas, nous obtenons des pas différents. 106 pas ont été récupérés en utilisant le filtre de Kalman, dont 39 sont de longueurs « normales » (supérieures à 50 cm) et 67 sont de « petites » longueurs (inférieures à 50 cm). 108 pas ont été récupéré en utilisant le filtre passe-bas dont 40 sont de longueurs « normales » et 68 sont de « petites » longueurs. Les erreurs entre les longueurs de pas fournies par le papier et par notre algorithme sont calculées pour chaque pas. L'erreur moyenne (en centimètres et en pourcentage) est calculée comme la moyenne sur toutes les erreurs, considérant toutes les données, ensuite considérant seulement les longueurs « normales » et finalement que les « petites » longueurs. Les résultats sont montrés à la Table 7.2. Les colonnes représentent la distinction entre les données, « Tous » signifie que tous les pas sont pris, « Normaux » signifie que tous les pas considérés comme « normaux » sont sélectionnés et « Petits » signifie que les petits pas sont analysés. Nous présentons les résultats pour le filtre de Kalman et pour le filtre passe-bas (noté « valeur filtre de Kalman | valeur filtre passe-bas » dans chaque case de la table).

Cette table nous montre qu'il y a une différence entre l'erreur moyenne obtenue avec le filtre de Kalman et avec le filtre passe-bas. En effet, pour toutes les longueurs de pas, l'erreur moyenne pour le filtre de Kalman est 5,3 cm, soit 11,8 %, et pour le filtre passe-bas nous obtenons une erreur moyenne de 6,9 cm, soit 15,6 %. La décomposition des données en pas normaux et petit pas nous permet de constater que l'erreur moyenne sur la série de données est principalement due aux erreurs sur les petits pas. L'erreur moyenne, pour le filtre de Kalman, pour les petits pas est de 4,7 cm, soit 12,8 % et pour les pas normaux, l'erreur est 6,4 cm, soit 10,0 %. L'erreur moyenne, avec le filtre passe-bas, pour les petits pas est de 6,5 cm, soit 18,0 % et pour les pas normaux, l'erreur est de 7,5 cm, soit 11,7 %.

Les boîtes à moustaches de la figure 7.4 permettent d'observer la variabilité des données. Les médianes des deux filtres sont identiques (représentées par la barre du milieu dans la boîte). Cependant, le filtre passe-bas ne lisse pas toujours les données correctement, ce qui provoque une erreur dans l'extraction des maxima locaux indiquée par les valeurs aberrantes de la figure. Les valeurs aberrantes sont représentées par les cercles en dehors de la boîte à moustache.

	Tous	Normaux	Petits
Erreur moyenne (cm)	5,3 6,9	6,4 7,5	4,7 6,5
Erreur moyenne (%)	11,8 15,6	10,0 11,7	12,8 18,0
Nombre de pas	106 108	39 40	67 68

TABLE 7.2 – Précision des longueurs de pas fournies par l'algorithme, obtenues avec un filtre de Kalman et avec un filtre passe-bas (pour tous les pas, pour les pas « normaux » et pour les petits pas).

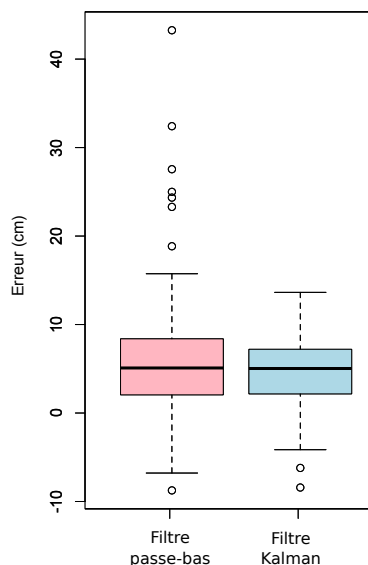


FIGURE 7.4 – Représentation des erreurs sur les longueurs de pas extraites à partir du filtre de Kalman et du filtre passe-bas.

7.1.3 Conclusion

Cette première expérience nous permet de conclure que la hauteur et l'angle de la caméra n'ont pas d'influence sur la précision des paramètres de la marche fournis par notre dispositif. Cette expérience nous permet également de constater qu'il est préférable de filtrer la trajectoire du centre de masse avec un filtre de Kalman. En effet, les longueurs de pas obtenues lorsque la trajectoire est filtrée par Kalman sont plus proches des valeurs réelles que les longueurs de pas obtenues lorsque la trajectoire est filtrée par un filtre passe-bas.

7.2 Évaluation globale du système

Les résultats obtenus dans la première évaluation permettent de constater qu'une expérience plus importante peut être réalisée et surtout permet de construire le nouveau protocole expérimental en prenant en compte les hypothèses que nous venons de tester.

7.2.1 Description de l'expérimentation

11 sujets sains ont intégré l'expérience, dont quatre femmes et sept hommes, âgés de 22 à 53 ans. L'expérience consiste à marcher sur un tapis actimétrique (le tapis de la marque GAITRite) de 5 mètres de longueur, comme montré à la figure 7.5. L'utilisation du tapis est une méthode plus coûteuse que la méthode « papier/encre », mais l'avantage est que nous pouvons traiter plus de données, car l'obtention des paramètres de la marche est automatisée. De plus, la récupération des longueurs de pas, dans la première évaluation, pouvait comporter des erreurs, car cette tâche était effectuée manuellement. L'erreur de mesure avec le tapis est 1,27 cm, représentant la distance entre chaque capteur. Les paramètres de la marche obtenus avec le tapis sont comparés aux valeurs fournies par l'algorithme de la caméra. Dans cette expérience, les trois paramètres de la marche sont comparés : les longueurs de pas, la cadence et la vitesse de la marche.

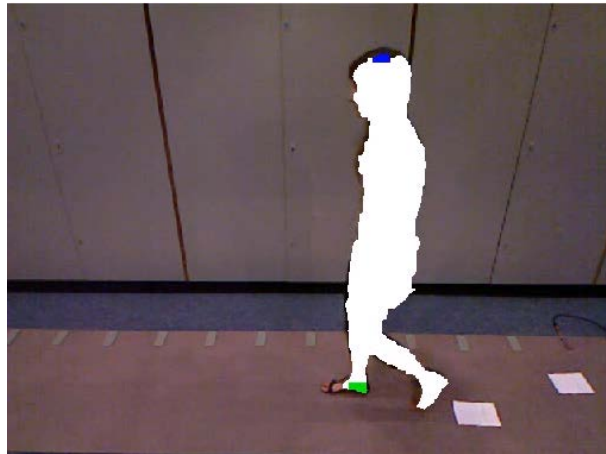


FIGURE 7.5 – Expérience comprenant un tapis actimétrique.

La première évaluation permet de structurer la nouvelle expérimentation. Notamment, elle nous a permis de constater qu'il n'y avait pas de différence significative concernant la précision des longueurs de pas liée à l'angle de la caméra (entre un angle de -27° et 0°). Donc, dans cette nouvelle expérience, nous considérons cette hypothèse de l'angle de la caméra comme avérée et c'est pourquoi nous ne testons pas différents angles, mais gardons le même pour toute l'expérience. La caméra est positionnée en hauteur avec un angle de -27° . De plus, la première évaluation a montré que le filtre de Kalman donnait une meilleure précision des longueurs de pas comparé au filtre passe-bas. Donc, nous analysons uniquement les paramètres extraits à partir du filtre de Kalman. Enfin, lorsque les longueurs de pas sur le papier ont été récupérées, le dernier pas n'était pas conservé, étant situé au bord de l'image à l'endroit où le sujet quitte le champ de vision de la caméra. Les derniers pas sont donc enlevés pour chaque séquence dans cette nouvelle expérience.

Comme dans la première évaluation, pour tester un large échantillon de données, les sujets marchent normalement puis font des petits pas. Trois nouvelles situations sont ajoutées :

- marcher avec une longue robe (allant jusqu'au pied et pour tous les sujets) pour savoir si les vêtements peuvent perturber l'extraction des paramètres de la marche ;
- marcher face à la caméra, toutes les autres situations étant réalisées avec la caméra positionnée perpendiculairement aux sujets ;
- marcher avec deux caméras observant la même scène pour savoir si l'interférence ne dégrade pas la précision des résultats.

Pour suivre une personne à son domicile nous avons en effet besoin de couvrir les différentes pièces avec plusieurs caméras. En plaçant dans une même pièce plusieurs caméras, il est probable qu'il y ait des recouvrements entre leurs champs de vision. Avec la dernière situation, nous voulons évaluer si dans cette zone commune les résultats sont dégradés, car la caméra RGB-D que nous utilisons est un capteur actif. Les 11 sujets réalisent les cinq situations trois fois pour permettre d'obtenir une base de données suffisantes en nombre de pas.

7.2.2 Résultats

Les valeurs des longueurs de pas, de la cadence et de la vitesse de marche sont mesurées avec le tapis et avec notre dispositif pour chaque séquence. Nous traitons 165 séquences, composées chacune de 2 à 5 pas, selon les sujets et la situation analysée. Les résultats sont présentés, dans

un premier temps, à travers une analyse de l'erreur sur chaque séquence, puis à travers une étude évaluant la moyenne des paramètres de la marche.

Résultats basés sur l'erreur de chaque séquence

Les erreurs entre les moyennes des paramètres de la marche, fournies par le tapis et notre algorithme, sont calculées pour chaque séquence. Les séquences sont ensuite regroupées par situations et nous présentons l'erreur moyenne calculée pour chaque situation. La table 7.3 montre les résultats. Dans chaque case de la table, le premier chiffre correspond soit aux erreurs moyennes des longueurs de pas exprimées en centimètres, soit aux erreurs moyennes de la cadence exprimée en nombre de pas par seconde, soit aux erreurs moyennes des vitesses en mètre par seconde. Puis le deuxième chiffre correspond à l'écart-type de ces données et le troisième à l'erreur en pourcentage.

	Normal	Petit	Jupe	2 caméras	Face
Long (cm σ %)	2,5 1,6 3,7	1,9 1,7 4,7	2,7 1,8 4,6	3,8 2,8 5,5	9,0 9,5 12,5
Cad (Pas/s σ %)	0,09 0,08 6,2	0,10 0,08 6,7	0,06 0,07 3,9	0,09 0,06 6,0	0,14 0,13 8,8
Vit (m/s σ %)	0,04 0,02 3,7	0,02 0,01 3,3	0,04 0,02 4,4	0,02 0,02 2,0	0,11 0,09 10,0
Nombre de pas	50	94	61	44	34

TABLE 7.3 – Erreurs moyennes des longueurs (Long), cadences (Cad) et vitesses (Vit) pour chaque situation.

Les erreurs moyennes des longueurs, concernant les situations « Normal », « Petit » et « Jupe », sont similaires avec une erreur moyenne de 3,7 % à 4,7 %, représentant une erreur de moins de 2,7 cm. L'erreur moyenne, pour la situation où il y a deux caméras, est un peu plus grande avec une erreur de 5,5 %. L'erreur moyenne des longueurs est extrêmement grande avec une valeur de 12,5 % pour la situation où la personne marche face à la caméra.

Les erreurs concernant les cadences sont plus grandes, à part pour la situation « Jupe » qui conserve une erreur d'environ 3,9 %. Les erreurs moyennes pour « Normal », « Petit » et « 2 caméras » sont entre 6 % et 6,7 %. L'erreur moyenne pour la situation « Face » est la plus grande, avec une erreur de 8,8 %.

Les erreurs commises pour les vitesses sont inférieurs aux erreurs des deux autres paramètres de la marche. Pour les situations « Normal », « Petit », « 2 caméras » et « Jupe », les erreurs moyenne vont de 2,0 % à 4,4 %. En revanche, l'erreur moyenne pour la situation « Face » est très grande, soit 10,0 %.

Résultats basés sur la moyenne de chaque paramètre de la marche

Une autre façon d'analyser les résultats est de calculer la moyenne, pour toutes les séquences d'une même situation, de chaque paramètres de la marche. Les résultats sont montrés à la table 7.4. Cette table représente toutes les moyennes et les écart-types (notés σ) calculés, pour chaque paramètre et chaque situation, à partir des mesures fournies par le tapis et par notre algorithme.

Pour les longueurs de pas, nous pouvons voir qu'il n'y a pas de grosses différences entre les moyennes obtenues avec le tapis et avec notre algorithme quelque soit la situation. L'écart maximal, entre le tapis et la caméra, est de 2,2 cm, obtenu pour la situation où les personnes portaient une jupe. Les situations « Normal », « 2 caméras » et « Petit » sont celles pour

Situation	Technique	Longueur(cm)	Cadence(Pas/s)	Vitesse(m/s)
		Moyenne σ	Moyenne σ	Moyenne σ
Pas normaux	Tapis	67,6 7,11	1,53 0,15	1,09 0,12
	Caméra	67,0 8,31	1,61 0,20	1,12 0,12
Petits pas	Tapis	38,4 6,82	1,44 0,19	0,57 0,11
	Caméra	38,3 9,09	1,51 0,24	0,59 0,12
Avec jupe	Tapis	57,1 8,72	1,46 0,31	0,92 0,14
	Caméra	59,3 9,93	1,49 0,31	0,96 0,13
Face caméra	Tapis	68,1 7,18	1,64 0,13	1,13 0,16
	Caméra	66,5 27,77	1,61 0,75	1,02 0,16
Deux caméras	Tapis	66,9 6,91	1,55 0,13	1,10 0,16
	Caméra	66,8 8,65	1,60 0,16	1,11 0,16

TABLE 7.4 – Comparaison des paramètres de la marche obtenus avec le tapis et avec notre algorithme utilisant la caméra.

lesquelles les écarts entre les moyennes du tapis et de la caméra sont les plus faibles. Pour la situation « Face », nous obtenons une erreur de 1,6 cm, mais avec une très grande variabilité. L'écart entre l'écart-type pour le tapis et pour la caméra est de 20 cm.

Pour la cadence, l'écart entre les moyennes du tapis et de la caméra n'excède pas une erreur de 0,08 pas par seconde.

Concernant la vitesse de marche, l'erreur est plus importante pour la situation « Face » que pour les autres situations. Pour la situation « Face », l'écart entre les moyennes est de 0,11 m/s alors que dans les autres situations, elle n'excède pas 0,04 m/s.

7.3 Discussion et conclusion

Une première évaluation nous a permis de constater que le système fournit des valeurs proches de la réalité. Nous en avons déduit qu'il était possible d'investir dans une expérimentation plus importante. Cette première évaluation nous a également permis de valider l'hypothèse que l'angle de la caméra n'a pas d'influence sur la précision des résultats. Aussi, elle nous a permis de remarquer que le lissage de la trajectoire du centre de masse avec le filtre de Kalman fournit une plus grande précision des longueurs de pas comparée à la précision obtenue avec le filtre passe-bas.

La seconde évaluation avait pour but de tester la précision des paramètres de la marche, obtenue avec la caméra, dans diverses situations pouvant être rencontrées au domicile. Ainsi lorsque deux caméras ont le même champ de vision, les résultats sont légèrement dégradés, comparés à la situation où la personne marche devant une seule caméra. Nous avons constaté que l'habillement n'avait, apparemment, pas d'influence sur la précision des résultats obtenus. A la section 4.2.2, nous avons décrit que les oscillations de la trajectoire du centre de masse sont dues au mouvement du centre de masse de la personne, amplifié par un artefact lié aux occlusions de profil du corps. La situation où la personne porte une jupe renforce l'idée de l'existence d'un artefact. En effet, les oscillations de la trajectoire du centre de masse, de la personne portant une jupe, sont plus grandes que celles où la personne porte un pantalon, comme montré à la figure 7.7. Le nombre de points détectés en bas du corps de la personne est encore plus grand quand elle marche vêtue d'une jupe, en position de double appui comme le montre la figure 7.6. Le

phénomène de l'artefact explique également pourquoi beaucoup d'erreurs sont commises quand la personne marche face à la caméra. Quand la caméra est face à la personne, l'artefact n'est plus présent, rendant la trajectoire du centre de masse plus lisse. L'atténuation de l'amplitude des oscillations est également due à la position de la caméra qui se trouve en hauteur, avec un angle de -27° comme nous l'avons déjà décrit à la section 4.2.2. La trajectoire étant plus lisse, des erreurs d'extraction des maxima locaux et, par conséquent, d'extraction des paramètres de la marche sont commises. Les résultats, de cette deuxième évaluation, nous montrent également que si l'on souhaite installer un système au domicile, analysant les paramètres de la marche, nous devons analyser les moyennes obtenues pour chaque paramètre, par exemple à la fin de la journée. Nous avons montré que les moyennes pour chaque paramètre sont justes et fidèles aux valeurs réelles et cela pour toutes les situations même celles où la personne marche face caméra.

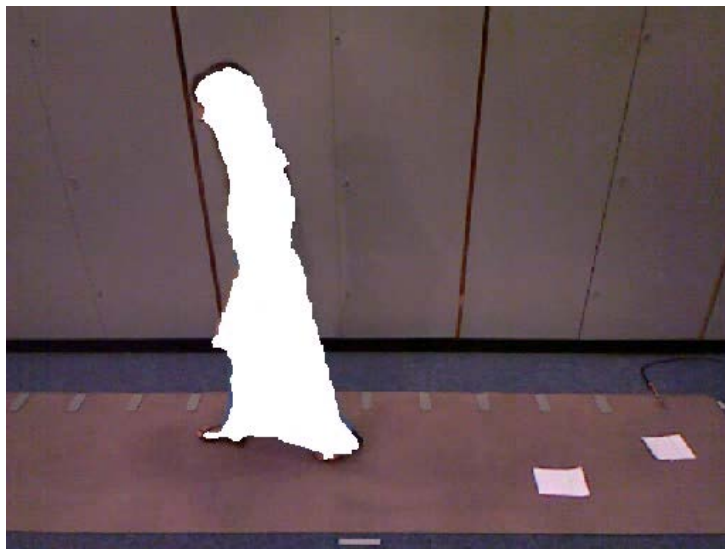
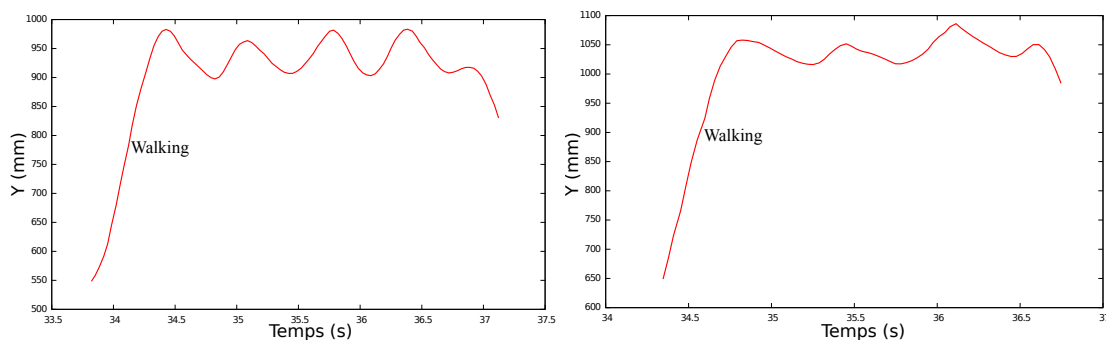


FIGURE 7.6 – En phase de double appui, de nombreux points mobiles sont détectés dans la partie basse de la silhouette quand le sujet porte une jupe.



(a) Trajectoire du centre de masse d'un sujet marchant avec une jupe.

(b) Trajectoire du centre de masse d'un sujet marchant avec un pantalon.

FIGURE 7.7 – Comparaison des oscillations de la trajectoire du centre de masse.

Chapitre 8

Expérimentations pour la reconnaissance d'activité

Sommaire

8.1	Cas de la personne entièrement visible	99
8.1.1	Description de l'expérimentation	99
8.1.2	Résultats	102
8.1.3	Conclusion	108
8.2	Cas de la personne partiellement visible	109
8.2.1	Description de l'expérimentation	109
8.2.2	Résultats	109
8.2.3	Conclusion	113

Nous présentons deux expérimentations qui ont pour but de tester la capacité des différents modèles de Markov cachés à reconnaître l'activité de la personne. Nous les comparons entre eux pour identifier celui reflétant au mieux la réalité. Cette comparaison est effectuée en prenant une unique base de données pour l'apprentissage contenant des situations où les personnes sont entièrement visibles. Dans la première expérimentation, nous testons les différents modèles de Markov cachés construits pour reconnaître l'activité sur des séquences où la personne est entièrement visible. Puis dans la seconde expérimentation, nous évaluons le comportement de chaque modèle construit, face aux situations où la personne n'est pas entièrement visible. Cette évaluation est nécessaire, car, même en plaçant les caméras de façon à maximiser la couverture de la pièce, il est fort probable que la personne ne soit pas entièrement visible dans certaines zones.

8.1 Cas de la personne entièrement visible

8.1.1 Description de l'expérimentation

Protocole expérimental

28 sujets, âgés de 20 à 53 ans, ont participé à l'expérience, dont 8 femmes et 20 hommes. Cette expérience s'est déroulée dans l'appartement laboratoire du centre de Nancy. Les sujets ont réalisés 8 situations filmés par une caméra RGB-D posée dans un angle de l'appartement. Le plan de l'appartement, avec la position de la caméra, est montré à la figure 8.1. La table représentée sur la figure, en face de la caméra, a été retirée pour permettre la réalisation des différentes situations. Chaque situation correspondait à une activité. Il s'agit des activités de la

vie quotidienne que nous avons définies précédemment, lors de la construction des modèles de Markov cachés au chapitre 6. Pour rappel, ces activités sont : s'accroupir, se pencher, s'asseoir, marcher, chuter, être allongé au sol, monter sur un obstacle et s'allonger sur un canapé. Plus précisément, la consigne donnée aux sujets était :

- pour la situation « Se pencher » : ramasser un stylo en pliant le moins possible les genoux ;
- pour la situation « S'accroupir » : ramasser un stylo en pliant les genoux, pour ainsi être plus proche du sol à la différence de la situation « Se pencher » ;
- pour la situation « S'asseoir » : s'asseoir sur une chaise ;
- pour la situation « Marcher » : marcher une fois en ligne droite et une fois en réalisant un demi-tour dans le champ de vision de la caméra. Ces deux séquences ont été regroupées dans l'unique situation « Marcher » ;
- pour les situations « Chuter » et « Allongé au sol » : chuter sur un matelas et rester allongé, enregistrés en une seule séquence qui a ensuite été découpée manuellement ;
- pour la situation « Monter sur un obstacle » : réaliser deux séquences, une en montant sur une chaise et une autre en montant sur un marche pied. La réalisation de deux séquences a permis d'apprendre la situation « Monté sur un obstacle » pour plusieurs tailles d'obstacle ;
- pour la situation « S'allonger sur un canapé » : s'allonger sur trois chaises mises côte à côte.

Les éléments tels que la chaise, le marche pied, le matelas, le stylo, étaient placés au milieu du champ de vision de la caméra. Avant d'atteindre ces éléments et de pouvoir réaliser la situation demandée, les personnes devaient d'abord marcher. Ensuite, après avoir réalisé ce qui était demandé, les sujets devaient remarcher, pour sortir du champ de vision de la caméra. Chaque situation est illustrée à la figure 8.2, montrant ainsi les étapes réalisées.

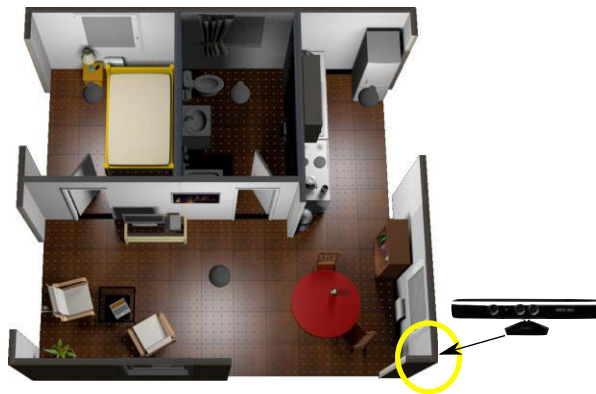


FIGURE 8.1 – Plan de l'appartement expérimental avec le positionnement de la caméra.

Par la suite, nous avons demandé à des sujets d'effectuer une situation où il devait marcher, se tenir debout à un endroit de la pièce, puis se déplacer pour rester à nouveau debout en répétant cela environ 5 fois. Ainsi, nous obtenons l'activité « Debout » à différentes distances de la caméra.

Base d'apprentissage et base de test

Les situations effectuées par les 28 sujets permettent de construire une base de données contenant 229 séquences. Chaque séquence a été segmentée et étiquetée manuellement. La base de données est utilisée pour apprendre les paramètres des modèles de Markov cachés et également pour tester la capacité des modèles à détecter correctement les activités réalisées.

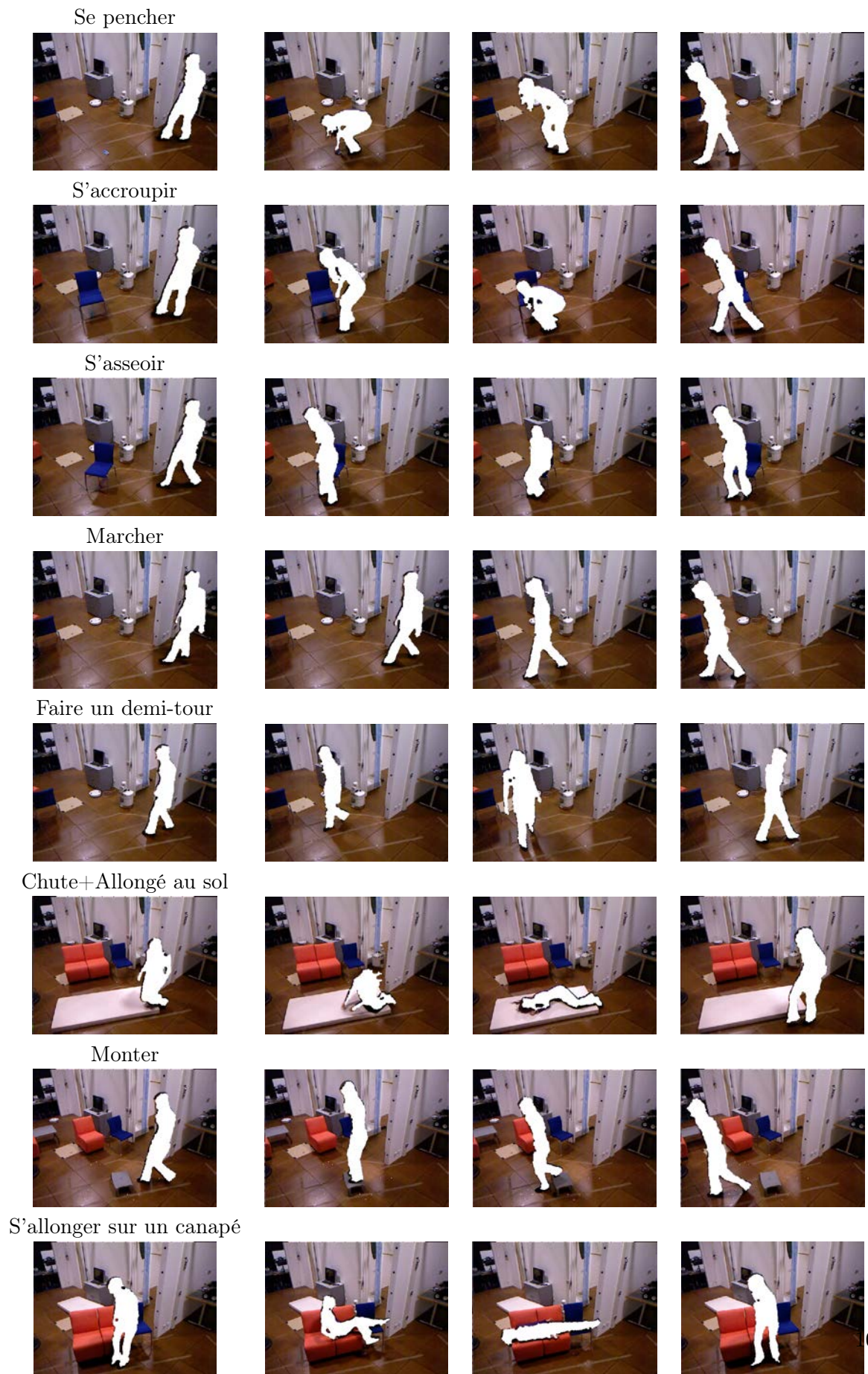


FIGURE 8.2 – Suivi des situations, effectuées par un sujet, à différents pas de temps.

185 séquences (soit 22 sujets) sont conservées pour l'apprentissage. Pour les modèles des sections 6.3.1 et 6.3.2, avec fonction d'observation seule et n'utilisant qu'un MMC, la moyenne de chaque paramètre (position verticale du centre de masse, sa vitesse verticale et la dispersion verticale du corps de la personne et pour certains modèles, la vitesse horizontale et le point maximum) et la covariance sont apprises à partir de cette base d'apprentissage. Elles sont également apprises pour chaque état du MMC, pour ainsi construire la fonction d'observation. Pour les modèles utilisant plusieurs MMC de la section 6.3.3, la matrice de transition, les probabilités initiales et la fonction d'observation (utilisant les moyennes et les covariances des trois ou cinq paramètres selon le modèle) sont apprises, pour chaque MMC, à partir de cette même base de données et ce grâce à l'algorithme de Baum-Welch.

Les 44 séquences restantes (correspondant à 6 sujets) sont utilisées pour tester le modèle. Nous fournissons à l'algorithme ces 44 séquences et nous observons si celui-ci détecte correctement l'activité réalisée. Les résultats sont montrés à la section suivante.

8.1.2 Résultats

Comme indiqué précédemment, 44 séquences sont données à chacun des modèles construits. Les algorithmes *Forward-Backward* et de Viterbi fournissent les classifications dans le cas des modèles à un seul MMC. La vraisemblance du modèle permet de classer les situations dans le cas de l'utilisation de plusieurs MMC.

Représentation des bonnes classifications et des fausses détections

Deux façons de représenter les résultats sont proposées pour évaluer les modèles, tout d'abord en observant les bonnes classifications puis en calculant le nombre de fausses détections.

Bonnes classifications Tout d'abord, le tableau 8.1 indique le nombre de bonnes classifications pour chaque activité selon le modèle utilisé. Les différents modèles sont exprimés dans ce tableau par des abréviations dont la signification est la suivante :

- « Modèle observations seules » correspond au modèle à 8 états avec la fonction d'observation seule de la section 6.3.1 ;
- « 1MMC 3 observations » représente le MMC à 8 états avec matrice de transition, définie manuellement, et 3 observations décrit à la section 6.3.2 ;
- « 1MMC 5 observations » représente le MMC à 9 états avec matrice de transition, définie manuellement, et 5 observations décrit à la section 6.3.2 ;
- « 8MMC 2 états » et « 8MMC 3 états » correspondent au système de 8 MMC à 2 ou 3 états et 3 observations décrit à la section 6.3.3 ;
- « 9MMC 2 états » et « 9MMC 3 états » correspondent au système de 9 MMC à 2 ou 3 états et 5 observations décrit à la section 6.3.3.

Les colonnes représentent les différentes activités (« Acc » : Accroupi, « All » : Allongé sur un Canapé, « Ass » : Assis, « Chu » : Chute, « Sol » : Allongé au sol, « Mar » : Marche, « Mon » : Monté sur un obstacle, « Pen » : Penché, « Dem » : Demi-tour). Nous séparons l'activité « Marche » et « Demi-tour » dans le tableau mais toutes les deux doivent être reconnues comme « Marche » par l'algorithme pour être correctes.

Dans chaque case du tableau, le premier chiffre (dans certain cas l'unique chiffre) correspond au nombre de bonnes classifications. Nous considérons comme « bonne classification » toutes les activités qui ont été réalisées dans une séquence et qui ont été détectées par l'algorithme avant la fin de l'activité. Par exemple, si une personne s'assoit et que l'algorithme détecte l'activité

« Assis » lorsque la personne se relève, nous considérons que l'algorithme n'a pas détecté la bonne activité, car il l'a détectée trop tard.

Dans certains cas, un deuxième chiffre est présent dans la case, correspondant au nombre de fois, inclus dans le premier chiffre, où l'algorithme donne la bonne activité en insérant, toutefois, une autre activité non réalisée dans la séquence. Ainsi, nous comptons comme juste cette séquence dans le premier chiffre mais nous indiquons son erreur par le deuxième chiffre. Nous considérons qu'il s'agit d'une classification correcte, car bien souvent l'insertion d'activité, non réalisée dans la séquence, se produit dans les phases de transition entre deux activités réalisées par la personne. Par exemple, lorsque la personne s'assoit, ce qui nous intéresse est que l'algorithme détecte correctement « Assis » quand la personne est effectivement assise et non pas quand elle passe de debout à assise. Pour expliquer la deuxième case du tableau, la situation « S'allonger sur un canapé », analysée par le modèle avec la fonction d'observation seule, contient « 5 | 2 » signifiant que 5 séquences ont été correctement identifiées et que, parmi ces 5 séquences, une autre activité a été indiquée 2 fois avant que la bonne activité soit finalement détectée.

La ligne « Total » du tableau indique le nombre total d'occurrences qui aurait dû être trouvé pour chaque activité. Par exemple, pour la deuxième case du modèle sans transition, l'algorithme a reconnu que la personne était dans la situation « Allongé Canapé » dans 5 séquences sur 6.

	Acc	All	Ass	Chu	Sol	Mar	Mon	Pen	Dem
Modèle observations seules	6 6	5 2	5 5	6 5	6	5	6	2	4 1
1MMC 3 observations	6	5	5	6	6	5	6	2	5
1MMC 5 observations	6	5	6	6	6	5	6	2	5
8MMC 2 états	6 3	4	6 1	6	4	5 1	4 1	4	5 1
8MMC 3 états	6	5	5	6	6	5	5	2	5
9MMC 2 états	1	5	6	6	1	5	4	4	4
9MMC 3 états	6	3	6	6	0	5	1	0	4
Total	6	6	6	6	6	5	6	4	5

TABLE 8.1 – Classification des activités par l'algorithme selon le modèle utilisé.

Le tableau 8.1 présente le nombre de fois où une séquence a été correctement détectée.

Fausse détections La deuxième représentation est de créer un tableau indiquant, quant à lui, les fausses détections. Les résultats sont présentés, pour chaque modèle, dans le tableau 8.2. Par exemple, pour le modèle sans transition, l'algorithme commet 2 fausses détections en détectant « Assis », 9 pour « Chuter » et 3 pour « Pencher ». Nous considérons une fausse détection lorsque l'activité réalisée n'est à aucun moment détectée dans la séquence ou l'est trop tard. Les activités « Chute » et « Allongé au sol » sont considérées comme des cas particuliers. A chaque fois que l'algorithme indique l'état « Chute » ou « Allongé Sol » alors que ces activités n'ont pas été réalisées, nous considérons qu'il commet une fausse détection, même si l'algorithme donne finalement la bonne activité. L'idée est de considérer la détection d'une chute (ou être allongé au sol) comme une information très importante et commettre une fausse détection de ces activités est plus grave que de produire une fausse détection des autres activités.

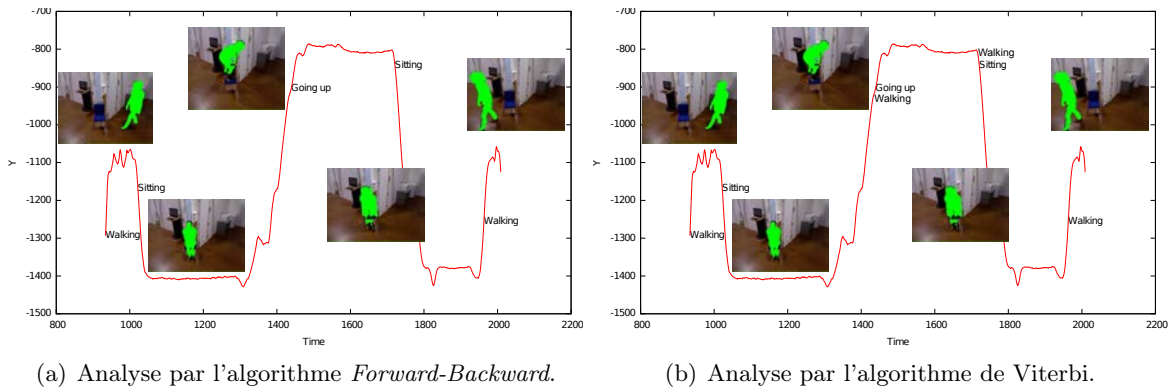
Comparaison entre les algorithmes *Forward-Backward* et de Viterbi

Les algorithmes *Forward-Backward* et de Viterbi s'utilisent dans le cas du modèle sans transition, du MMC à 8 et 9 états avec matrice de transition. Dans les tableaux 8.1 et 8.2, les résultats

	Acc	All	Ass	Chu	Sol	Mar	Mon	Pen
Modèle observations seules	0	0	2	9	0	0	0	3
1MMC 3 observations	1	0	2	0	0	0	0	1
1MMC 5 observations	1	0	1	1	0	0	0	0
8MMC 2 états	4	0	0	3	0	0	0	3
8MMC 3 états	2	0	2	0	0	0	0	1
9MMC 2 états	0	0	1	3	0	2	0	6
9MMC 3 états	13	0	0	3	0	0	0	0

TABLE 8.2 – Fausses détections effectuées par l'algorithme selon le modèle utilisé.

des trois premières lignes concernent l'algorithme *Forward-Backward*. La distinction entre les deux algorithmes n'est pas présentée, car peu de différences existent au niveau des résultats. La seule différence (mais elle n'est de toute façon pas prise en compte dans les tableaux) se situe au niveau des « états de transition » comme montré à la figure 8.3. Celle-ci présente les résultats de l'analyse effectuée pour les deux algorithmes, avec le MMC à 8 états et 3 observations, à partir d'une même séquence. Dans cette séquence, le sujet devait s'asseoir et se lever sur une chaise sans marcher entre les deux situations. *Forward-Backward* infère correctement les bonnes activités comme le montre la figure 8.3(a). L'algorithme de Viterbi de la figure 8.3(b) fournit plus d'états. Pour passer de « Assis » à « Monté sur un obstacle », le sujet ne marche pas, or l'algorithme de Viterbi infère l'état « Marche ». Cette décision de l'algorithme de Viterbi n'est pas correcte, mais il a l'avantage de coller aux contraintes fournies par le MMC. Dans l'approche à un unique MMC, où la matrice de transition est définie manuellement, l'état « Marche » est défini comme précédent et suivant l'état « Assis ». Donc la figure 8.3(b) montre que l'algorithme de Viterbi passe par l'état « Marche » avant et après l'état « Monté sur un obstacle » respectant ainsi le modèle défini, même si le sujet ne marche pas.

FIGURE 8.3 – Différence dans l'analyse faite par les algorithmes *Forward-Backward* et de Viterbi, pour une séquence où la personne s'assoit puis monte sur une chaise.

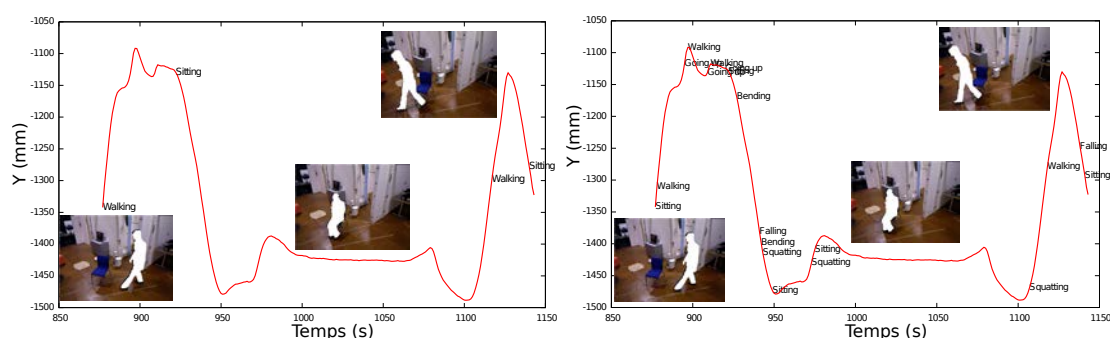
Pour la suite, les résultats seront donnés uniquement pour l'algorithme *Forward-Backward*, car nous préférons conserver l'algorithme qui colle aux situations réalisées plutôt que de conserver l'algorithme qui colle au modèle défini manuellement.

Modèle avec fonction d'observation seule

D'après le tableau 8.1, le modèle sans matrice de transition, détecte toutes les activités correspondant à « Accroupi », « Chute », « Allongé au sol », « Marche » et « Monté sur un obstacle ». Pour l'activité « Allongé Canapé », « Assis » et « Demi-tour » l'algorithme manque une détection.

Malgré une bonne détection des chutes, la performance de ce modèle est atténuée par le nombre important de fausses détections concernant les chutes comme montré à la Table 8.2. Pour chaque état nécessitant une descente du corps de la personne (par exemple « Assis », « Accroupi »), l'algorithme détecte une chute puis rectifie en donnant le bon état.

Les bonnes classifications sont rarement données directement comme on peut le voir par la présence du deuxième chiffre, dans les cases de la première ligne du Tableau 8.1, indiquant que la décision est bruitée. C'est-à-dire que l'algorithme, avant de donner le bon état, donne d'autres états non atteints dans la réalité. Cela peut être constaté par la figure 8.4(b) représentant la trajectoire du centre de masse sur l'axe vertical et le déroulement des décisions de l'activité dans le temps, par le modèle avec fonction d'observation seule. La figure représente le centre de masse d'une personne s'asseyant et montre que l'algorithme hésite à plusieurs reprises avant de donner la bonne activité réalisée « Assis ».



(a) Séquence analysée par le MMC à trois observations. (b) Séquence analysée par le modèle sans transition.

FIGURE 8.4 – Comparaison d'une séquence où la personne s'assoit analysée par deux modèles.

MMC à 8 états et 3 observations

Les résultats pour le MMC avec une matrice de transition montrent que les activités « Accroupi », « Chute », « Allongé au sol », « Marche » (pour les séquences où la personne marche et où elle fait un demi-tour) et « Monté sur un obstacle » sont toutes reconnues par le modèle. L'activité « Penché » est l'activité la moins bien reconnue, seulement deux fois sur quatre. Cette activité est remplacée par l'état « Assis ». Cette information est renseignée par le tableau des fausses détections 8.2. Le modèle commet une erreur pour l'activité où la personne s'allonge sur un canapé en désignant cette activité comme étant l'état « Accroupi ». Cette erreur vient du fait que la personne ne s'allonge pas complètement sur le canapé mais s'appuie sur son coude comme le montre la figure 8.5 provoquant une situation différente de ce qui a été appris par le modèle. Le modèle ne reconnaît pas non plus une situation où la personne est assise. En réalité, l'algorithme indique correctement l'état « Assis » lors de la descente et la remontée de la personne mais lorsque celle-ci est assise, le modèle reconnaît cette situation comme étant l'état « Penché », comme illustré à la figure 8.6. Sur cette figure, nous pouvons comprendre l'erreur

qui a été commise du fait que la personne en s'asseyant a remonté sa jambe sur son autre jambe. Ainsi, la trajectoire du centre de masse est remontée correspondant, pour le modèle, à l'état « Penché ».



FIGURE 8.5 – Sujet s'allongeant différemment de ce qui a été appris dans la base de donnée.

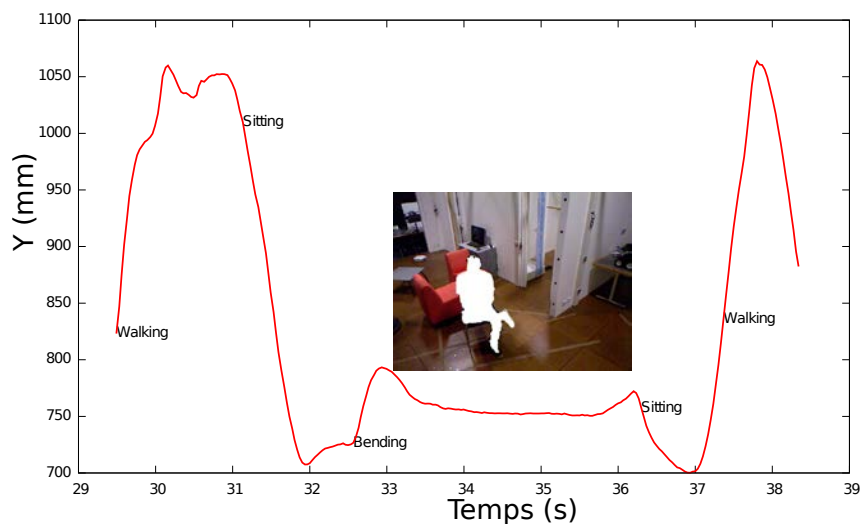


FIGURE 8.6 – Sujet s'asseyant différemment de ce qui a été appris dans la base de donnée.

Ce modèle détecte les bonnes activités de façon similaire au modèle sans matrice de transition. De plus, il présente l'avantage de faire moins de fausses détections comme le montre le tableau 8.2 et aucune fausse détection concernant les chutes. Ce modèle diffère également du précédent par le fait qu'il donne la bonne activité sans hésiter avec d'autres états (ceci est démontré par l'absence du deuxième chiffre dans les cases du tableau 8.1). Cette différence entre les deux modèles peut être observée à la figure 8.4. Les deux figures représentent la trajectoire du centre de masse sur l'axe vertical pour une personne marchant puis s'asseyant et enfin marchant pour sortir du champ de vision de la caméra. Ces figures indiquent les détections, par l'algorithme, des états dans lesquels est passée la personne. La figure 8.4(b) montre que le modèle sans transition est plus bruité que le modèle avec matrice de transition représenté à la figure 8.4(a). Selon nos critères, la classification pour le modèle sans transition est correcte, puisqu'au final l'algorithme

indique l'état « Assis », mais la décision apparaît plus lentement comparé au modèle avec des transitions, qui lui n'hésite pas sur l'activité réalisée.

MMC à 9 états et 5 observations

Les résultats pour le MMC à 9 états montrent que la classification des activités est semblable au modèle précédent. Dans ce modèle, les situations « Assis » sont toutes reconnues. De plus, toutes les activités sont reconnues directement, sans hésitation avec une autre activité comme le montre l'absence du deuxième chiffre dans les cases du tableau 8.1.

Le modèle à 9 états fait moins d'erreur au niveau des fausses détections que le précédent modèle à 8 états. Par contre, l'inconvénient est qu'il commet une fausse détection importante en donnant l'activité « Chute » à la place de l'état « Penché ».

Système de 8 MMC à 2 états et 3 observations

La reconnaissance d'activité est similaire aux autres modèles avec néanmoins les activités « Monté » et « Allongé Canapé » qui sont moins bien reconnues. Toutes les activités, où la personne se penche, sont ici reconnues. Les séquences où la personne « Chute » sont toutes reconnues, par contre l'activité « Allongé au sol » suivant la chute n'est pas toujours reconnue. La matrice de transition forçant le passage de « Chute » à « Allongé au sol » n'est plus présente dans ce modèle, cela permet de détecter l'état « Chute » sans forcément être suivi de l'état « Allongé au sol ». Remarquons que, dans certain cas les sujets se levaient directement après leur chute ne passant pas forcément par la situation « Allongé au sol ». Il y a un peu d'hésitation sur l'activité réalisée, la bonne classification ne se faisant pas toujours directement (comme on peut le voir avec la présence du deuxième chiffre du Tableau 8.1).

Peu de fausses détections sont présentes ici, mais l'inconvénient est qu'il y a trois détections de « Chute » non pertinentes.

Système de 8 MMC à 3 états et 3 observations

Le nouveau modèle classe correctement les activités. L'état « Penché » reste celui qui est le moins bien distingué. Les décisions de l'algorithme, concernant l'état dans lequel se trouve la personne, sont directes comme on peut le voir par l'absence d'un deuxième chiffre dans la ligne « 8MMC 3 états » du Tableau 8.1.

La différence avec le précédent modèle se trouve également au niveau des fausses détections. L'algorithme ne commet aucune fausse détection concernant les chutes. De plus, peu d'erreurs sont commises en règle générale par ce modèle.

Système de 9 MMC à 2 états et 5 observations

Le tableau 8.1 montre que ce modèle classe correctement les activités sauf pour deux états, « Accroupi » et « Allongé Sol ». Les situations « Allongé Sol » ne sont pas reconnues mais cela ne gêne pas la détection de l'état « Chute ».

Ce modèle commet peu de fausses détections. Les erreurs proviennent surtout de l'état « Accroupi » qui n'est pas détecté. L'algorithme remplace toutes les situations où la personne s'accroupit par l'état « Chute » suivi de l'état « Penché ».

Système de 9 MMC à 3 états et 5 observations

Ce modèle commet de nombreuses erreurs de classification notamment pour les états « Allongé Canapé », « Monté », « Allongé Sol » et « Penché ». Les états « Allongé Sol » et « Penché » ne sont d'ailleurs jamais reconnus. L'état « Penché » est confondu avec l'état « Accroupi ». L'état « Allongé Sol » n'est jamais reconnu mais cela n'empêche pas la chute d'être quant à elle détectée.

Le modèle infère l'état « Accroupi » très souvent puisque 13 fausses détections ont été répertoriées pour cet état. Le modèle commet également trois fausses détections concernant l'état « Chute ».

Résultats pour l'activité debout

Nous avons demandé à une personne de se tenir debout dans le but de savoir ce qu'infèrent les différents modèles. Le modèle à 9 états 5 observations et les systèmes de 9 MMC ont appris à différencier l'activité « Debout », donc il devrait la reconnaître. En revanche, les modèles à 8 états et les systèmes de 8 MMC n'ont pas eu, dans leur base d'apprentissage, une telle activité. Nous voulons ainsi connaître l'activité inférée, en pensant qu'ils vont la classer comme étant l'activité « Marche ». Les résultats sont montrés à la table 8.3.

	Debout
Modèle observations seules	Marche
1MMC 3 observations	Marche
1MMC 5 observations	Debout
8MMC 2 états	Marche
8MMC 3 états	Marche
9MMC 2 états	Assis
9MMC 3 états	Accroupi

TABLE 8.3 – Classification de l'activité « Debout » selon le modèle utilisé.

Les résultats montrent que seul le modèle à 9 états et 5 observations détecte l'état « Debout ». La plupart des autres modèles infèrent l'état « Marche ».

8.1.3 Conclusion

L'expérimentation nous a permis de comparer les différents modèles construits sur une même base de test. Les résultats nous permettent de savoir quels sont les modèles les plus justes. La classification réalisée par le modèle sans matrice de transition, après de nombreuses hésitations entre différents états, donne finalement de bons résultats. Mais ce modèle ne peut pas être conservé, car il fournit beaucoup de fausses détections concernant les chutes. Le modèle avec transitions à 8 états et 3 observations produit peu de fausses détections et surtout aucune concernant les chutes. La probabilité de transition d'aller dans l'état « Chute » a été définie manuellement comme très faible. Ainsi, l'observation doit être vraiment « forte » pour passer dans l'état « Chute » avec ce modèle. Le modèle réalise peu d'erreurs de classification. L'état le moins bien détecté est « Penché » en omettant 2 situations sur 4. Toutes les chutes sont détectées. Le modèle à 9 états et deux observations en plus des modèles précédents n'apporte pas de meilleurs résultats de classification. Au contraire, il commet une fausse détection concernant l'état « Chute », ce que ne faisait pas le modèle précédent. L'idée en combinant plusieurs modèles de Markov cachés est de

tester des modèles sans connaissance *a priori*. Le système de 8 MMC, appris par Baum-Welch, à 3 états et 3 observations s'avère le modèle s'approchant au mieux de la réalité. Ce modèle donne de bons résultats au niveau de la classification et c'est le seul parmi les systèmes de MMC à ne faire aucune fausse détection concernant les chutes.

Pour développer un algorithme capable de détecter les activités de la personne au cours de la journée et de détecter ses chutes, nous avons besoin d'un modèle classifiant les bonnes activités et surtout un modèle effectuant peu de fausses détections. L'algorithme à un modèle de Markov caché à 8 états et celui utilisant le système de 8 MMC à trois états détectent toutes les chutes sans produire de fausse détection et classifient correctement les autres activités, ce qui correspond à nos attentes. Ces deux modèles sont équivalents au niveau des résultats, mais on préfère un modèle ayant appris par lui-même les différentes probabilités. En effet, le système de 8 MMC, n'étant pas contraint au niveau de la dynamique des situations, est plus général et correspond alors mieux aux situations non prévues qu'un modèle où nous avons inséré de la connaissance.

Ces résultats nous montrent que les modèles à trois observations sont tout autant, voire plus justes que les modèles à cinq observations. Bien que nous ayons ajouté un état et qu'il est donc possible de distinguer l'état « Debout » avec 5 observations, la qualité de classification s'avère dégradée pour certains modèles. Donc, en ajoutant des états et des observations, nous ajoutons de la complexité et perdons au niveau précision des résultats.

8.2 Cas de la personne partiellement visible

8.2.1 Description de l'expérimentation

Nous avons demandé à une personne d'effectuer des activités en étant cachée par des objets. Ces activités correspondent à celles utilisées pour construire les MMC. Les consignes pour réaliser ces neuf situations sont les suivantes :

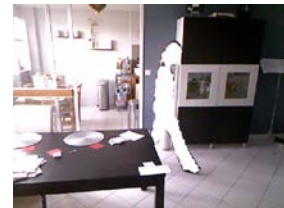
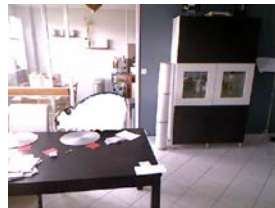
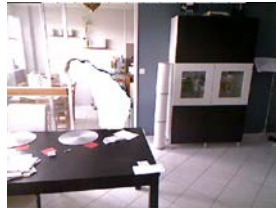
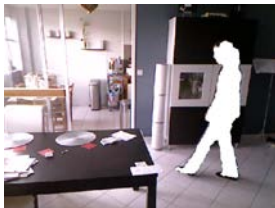
- la personne se penche derrière une table ;
- la personne s'accroupit derrière une table ;
- la personne s'assoit sur une chaise placée derrière une table ;
- la personne marche derrière trois chaises posées l'une à côté de l'autre ;
- la personne marche et réalise plusieurs demi-tours derrière les trois chaises ;
- la personne chute et s'allonge au sol derrière une table ;
- la personne s'allonge sur un canapé en posant un drap sur elle ;
- la personne monte sur une chaise placée derrière une table ;
- la personne reste debout derrière une table.

Chaque situation est illustrée à la figure 8.7, montrant ainsi les étapes réalisées par le sujet à chaque séquence. La table correspond à une table à manger d'environ 80 cm de hauteur. Lorsque la personne chute, elle n'est donc plus visible. Dans le cas des autres situations, une partie du corps de la personne reste toujours visible.

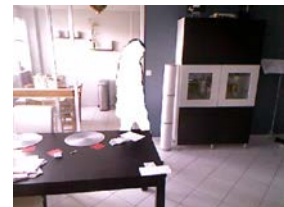
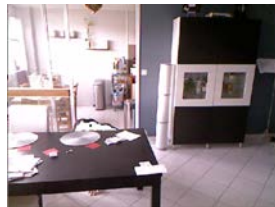
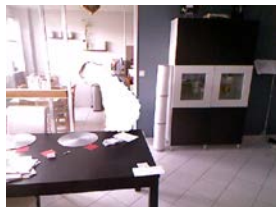
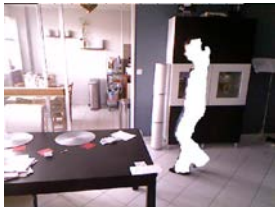
8.2.2 Résultats

Chaque situation, que devait réaliser le sujet, est analysée par les différents modèles construits. Nous avons regroupé dans la table 8.4 les activités inférées par chaque modèle à chaque situation. Les colonnes représentent les différentes activités réalisées par la personne (« Acc » : Accroupi, « Ass » : Assis, « Chu » : Chute, « Mar » : Marche, « Mo » : Monté sur un obstacle, « Pen » : Penché, « Dem » : Demi-tour, « All » : Allongé sur un canapé, « De » : Debout). Chaque case

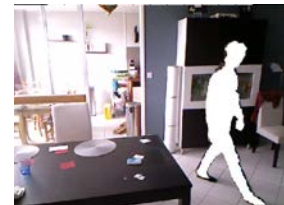
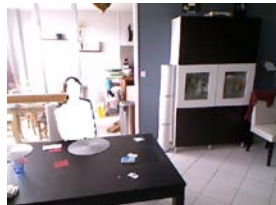
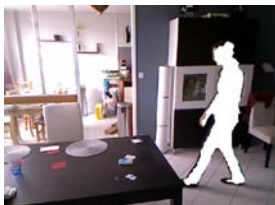
Se pencher



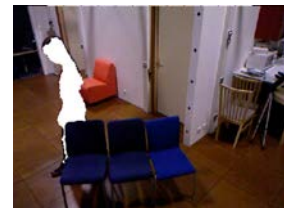
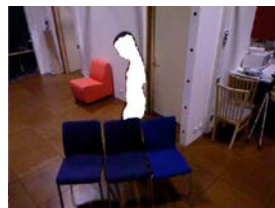
S'accroupir



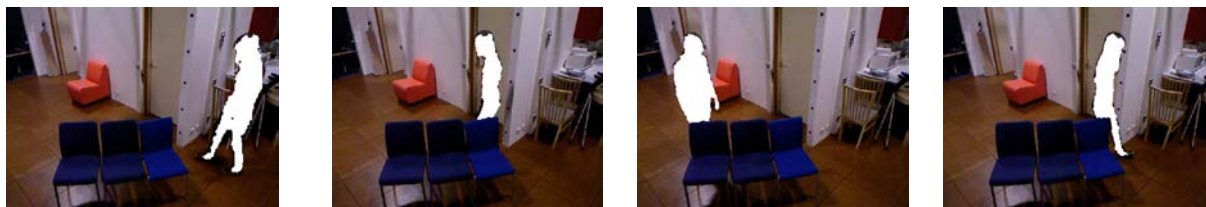
S'asseoir



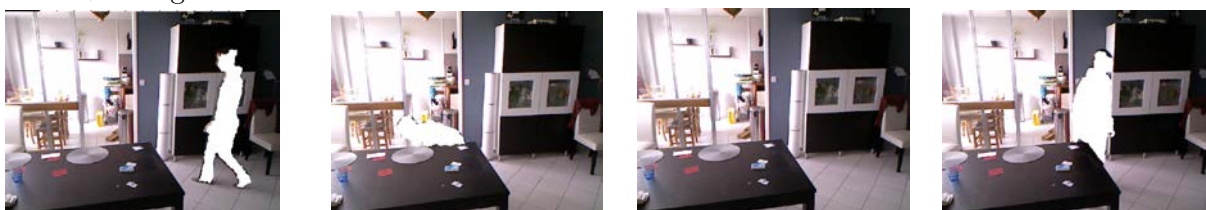
Marcher



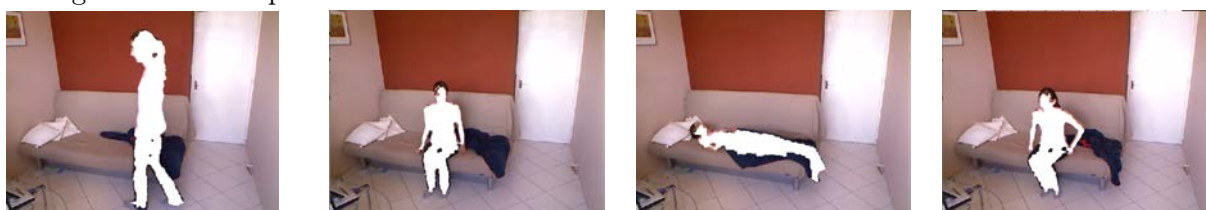
Faire des demi-tours



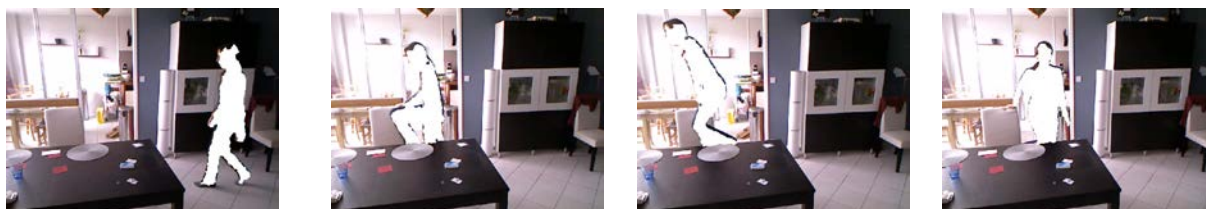
Chute+Allongé au sol



S'allonger sur un canapé



Monter



Debout

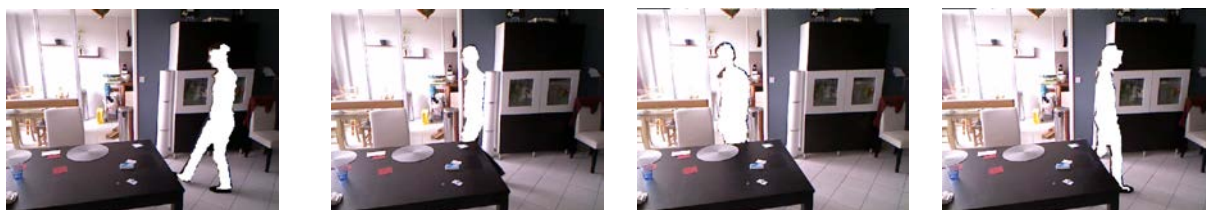


FIGURE 8.7 – Suivi des situations où le sujet n'est pas entièrement visible à différents pas de temps.

indique l'activité que le modèle a inféré. Les activités inscrites en vert correspondent aux bonnes classifications.

Le précédent tableau 8.1 indiquait les bonnes classifications dans le cadre de situations semblables à celles contenues dans la base de données. Dans cette nouvelle table 8.4, les erreurs de classifications sont plus fréquentes, car il s'agit de situations nouvelles, non apprises, où la personne est à moitié visible par la caméra. Le tableau nous informe que l'activité « Monté » est l'activité la mieux détectée par presque tous les modèles. La situation « S'allonger sur un canapé », sous un drap, est particulière, car sa réalisation entraîne la modification du fond appris. Lorsque la personne s'allonge, elle prend le drap pour le mettre sur elle. Ce qui a pour conséquence de cacher son corps et de sortir le drap du fond appris. Comme nous pouvons le voir sur les images de la figure 8.7, le drap est détecté comme mobile. La plupart des modèles reconnaissent l'activité « Assis », qui précède et suit l'activité « Allongé Canapé ». Cependant l'activité « Allongé Canapé » en elle-même, lorsqu'elle est réalisée sous un drap, n'est pas reconnue.

Le modèle commettant le moins d'erreurs de classification est le modèle à 9 états et 5 observations. De plus, dans la situation où la personne a réalisé plusieurs demi-tours derrière des chaises, ce modèle est le seul à indiquer les demi-tours comme le montre la figure 8.8. L'activité « Demi-tour » n'est pas contenue dans la base de donnée, mais le modèle à 9 états et 5 observations infère l'état « Debout » lorsque la personne tourne sur elle-même pour réaliser le demi-tour, puis infère l'état « Marche » entre chaque deux demi-tours.

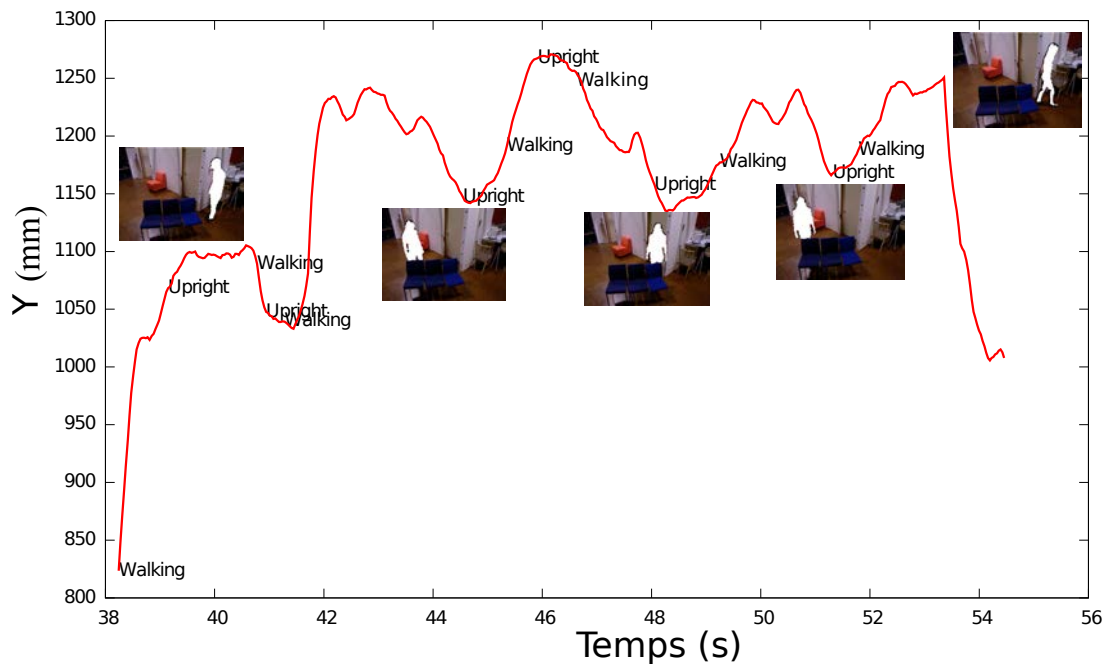


FIGURE 8.8 – Résultat de l'analyse avec le modèle à 9 états et 5 observations, pour une personne réalisant des demi-tours.

Pour l'ensemble des modèles, nous pouvons remarquer qu'il y a confusion entre les états « Accroupi », « Penché » et « Chute ». Ce problème peut être résolu en ajoutant la notion de temps. Par exemple, pour l'état « Accroupi » et « Penché » la personne reste très peu de temps dans ces activités alors que l'activité « Chute » correspond à une situation où la personne n'est plus visible durant une certaine période.

	Acc	Ass	Chu	Mar	Mo	Pen
Modèle observations seules	Acc	Pen	Acc	Mon + Ma	Mo	Pe
1MMC 3 observations	Chu + Acc	Acc	Chu	Ma	Mo	Acc
1MMC 5 observations	Chu + Acc	Ass	Chu	Ma	Mo	Acc
8MMC 2 états	Acc	Acc	Acc	Acc	Mo	Acc
8MMC 3 états	Acc	Acc	Acc	Acc	Mo	Acc
9MMC 2 états	Acc	Ass	Ass	Ma	Mo	Acc
9MMC 3 états	Acc	Acc	Acc	Acc	Acc	Acc

	Dem	All	De
Modèle observations seules	Mo + Acc + Ma	Acc	Ma
1MMC 3 observations	Ma	Acc	Acc + Ma
1MMC 5 observations	De + Ma	Acc	De
8MMC 2 états	Acc	Acc	Acc
8MMC 3 états	Acc	Acc	Acc
9MMC 2 états	Ma	Acc	Ass + Ma
9MMC 3 états	Acc	Acc	Acc

TABLE 8.4 – Détection d’activité quand la personne n’est pas entièrement visible pour différents modèles.

8.2.3 Conclusion

Avec la première expérimentation pour la reconnaissance d’activité d’une personne entièrement visible, nous avons conclu que le modèle à 8 états 3 observations et le système de 8 MMC étaient ceux commettant le moins d’erreurs de classification et peu de fausses détections. Le modèle à 9 états et 5 observations avait l’avantage de distinguer l’activité « Debout » de « Marche ». Mais ce modèle ajoutait de la complexité, en ajoutant deux observations de plus que ceux à 8 états, sans pour autant donner de meilleurs résultats. Avec la deuxième expérimentation pour la reconnaissance d’activité d’une personne partiellement visible, nous avons montré que le modèle à 9 états et 5 observations est celui donnant les meilleurs résultats pour traiter les situations d’occlusion de la personne non apprises dans la base de données. Pour le domicile, nous pensons que ce modèle fournira les meilleures classifications.

Chapitre 9

Tests en situations réelles

Sommaire

9.1	Test de l'algorithme d'analyse de la marche	115
9.2	Test de l'algorithme de détection d'activité	116
9.3	Conclusion	118

Une limite des expérimentations, décrites dans les deux chapitres précédents, est qu'elles ont été réalisées en laboratoire avec des sujets ne représentant pas la catégorie à laquelle cette thèse s'adresse. Ces expérimentations doivent être effectuées avec des sujets appropriés pour savoir si les résultats restent les mêmes. Dans ce chapitre, nous présentons des résultats de quelques tests effectués sur des personnes âgées.

9.1 Test de l'algorithme d'analyse de la marche

6 personnes âgées ont réalisé quelques séquences filmées par notre dispositif. Ces enregistrements ont été réalisés dans des milieux réels en dehors de l'appartement expérimental.

Tout d'abord, nous avons demandé à deux personnes, à l'hôpital, de marcher devant notre système. Ces deux patientes, ayant des problèmes liés à la marche et à l'équilibre, sont venues pour réaliser un bilan médical avec un médecin. Les consultations durent une heure, pendant laquelle de nombreux tests cliniques sont effectués pour constituer un bilan de la fragilité de la personne. Un des tests cliniques réalisé était de marcher sur 10 mètres. Durant ce test, nous avons filmé ces deux personnes. La figure 9.1 montre la marche de l'une des deux personnes et la trajectoire de son centre de masse, sur l'axe vertical, avec les longueurs de pas détectées. Pour cette expérience, nous n'avons pas de système de référence. Nous ne savons donc pas si notre système fournit des paramètres de la marche justes par rapport à la réalité. Nous pouvons, uniquement, effectuer une analyse qualitative en observant que le système détecte, comme dans nos précédentes expérimentations effectuées sur les sujets jeunes, des oscillations et est capable d'extraire les paramètres de la marche. De plus, nous pouvons remarquer que les longueurs de pas détectées (68 cm, 53 cm et 55 cm) sont des mesures plausibles et que leur nombre est juste.

Nous avons également filmé quatre personnes de plus de 80 ans, chez elles. Trois personnes sur quatre ont déjà chuté. Nous avons demandé aux personnes de marcher perpendiculairement à la caméra que nous avons posée dans le salon, sur des meubles hauts. Dans les quatre cas, les oscillations du centre de masse sont présentes comme nous pouvons le voir sur la figure 9.2. Cette figure présente les trajectoires du centre de masse, sur l'axe vertical, des quatre personnes. Les maxima des oscillations étant perceptibles, l'algorithme peut en déduire les paramètres de

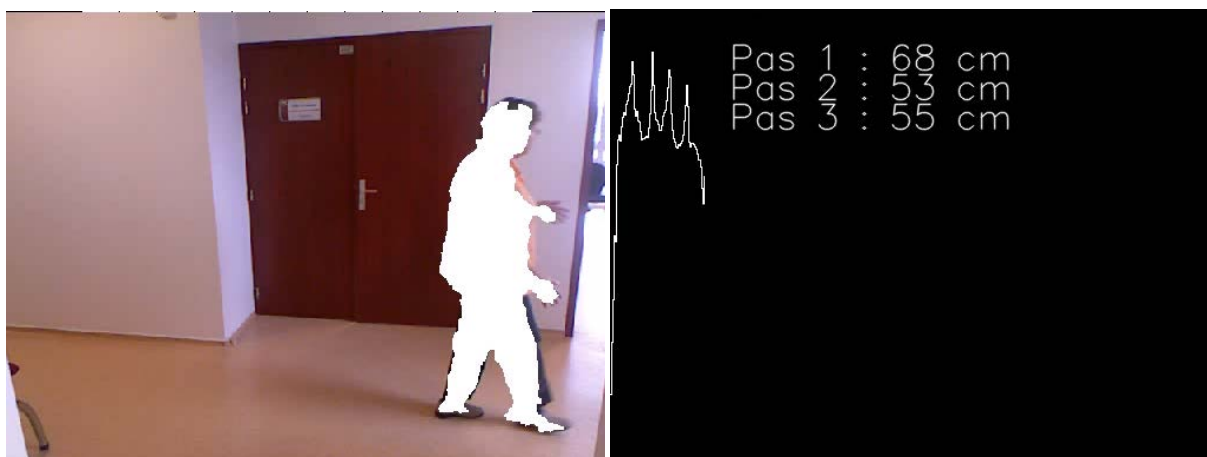


FIGURE 9.1 – Utilisation de notre système avec une patiente, lors de sa consultation à l'hôpital.

la marche. Nous n'avons, toutefois, pas de méthode de référence pour savoir si la localisation des oscillations permet de fournir les paramètres de la marche proche de la réalité. L'analyse qualitative nous permet de constater que la trajectoire du centre de masse la plus régulière, parmi les figures 9.2, appartient à la personne âgée n'ayant jamais chuté 9.2(a). Parmi les quatre personnes, une des personnes a d'importants problèmes d'équilibre. La personne a déjà chuté plusieurs fois. La trajectoire de son centre de masse, sur l'axe vertical, est représentée à la figure 9.2(d). Les oscillations sont moins régulières que ce que nous pouvons voir avec des personnes jeunes, mais la trajectoire reste sinusoïdale. Il est donc possible d'obtenir les paramètres de la marche. Une autre personne, de 84 ans ayant récemment chuté, a marché devant notre système. La trajectoire de son centre de masse, représentée à la figure 9.2(c), oscille peu. L'espacement des maxima indique une cadence plus lente que pour la personne représentée à la figure 9.2(a). Quand la personne a marché, nous avons pu voir qu'elle restait très droite. Pour finir, un des sujets a une jambe artificielle et a donc marché avec 2 cannes. Quand il marche, la jambe artificielle est peu soulevée du sol. La représentation de la trajectoire de son centre de masse est montré à la figure 9.2(b).

Nous pouvons voir que, malgré les différentes façons de marcher, différentes de ce que nous pouvions avoir avec les sujets jeunes, le système fournit le même type de trajectoires, même si elles sont moins régulières avec les personnes âgées qu'avec les plus jeunes.

9.2 Test de l'algorithme de détection d'activité

L'algorithme de détection d'activité fournit de bons résultats au niveau de la classification, comme nous l'avons vu au chapitre 8, sur des personnes jeunes. Nous avons testé l'algorithme pour quelques activités réalisées par des personnes âgées. Pour les précédentes situations où la personne marchait devant la caméra à leur domicile ou à l'hôpital, l'algorithme de détection d'activité a correctement classifié les personnes en train de marcher. Comme nous pouvons le voir à la figure 9.3, pour une personne, marchant à son domicile, l'algorithme indique qu'elle marche.

Nous avons également testé la capacité de l'algorithme à détecter d'autres activités. Nous avons demandé à un des sujets de s'asseoir, de se pencher et de monter sur un marche pied. La figure 9.4 montre le résultat pour chacune des activités réalisées par le sujet. Chaque figure représente la trajectoire du centre de masse de la personne, sur l'axe vertical, avec l'évaluation par

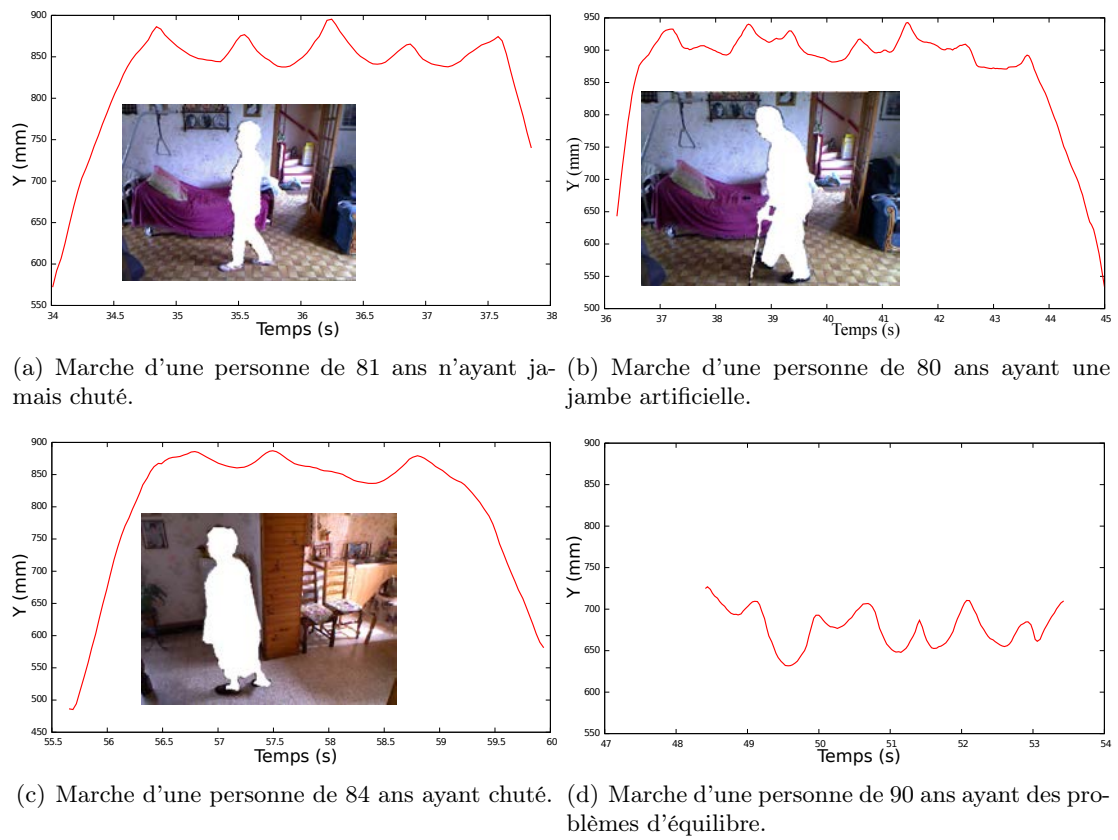


FIGURE 9.2 – Représentation du centre de masse, sur l'axe vertical, de quatre personnes âgées.

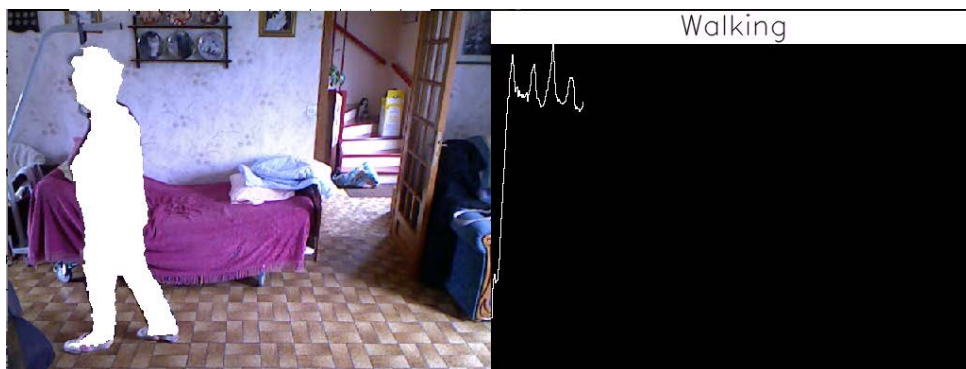
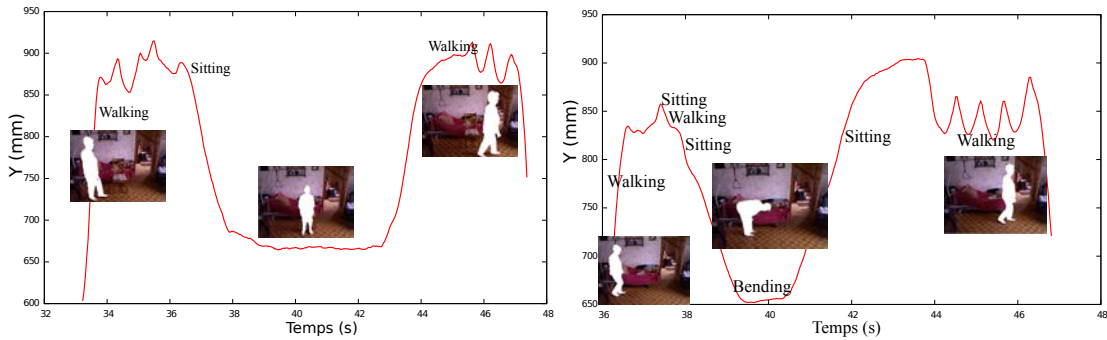
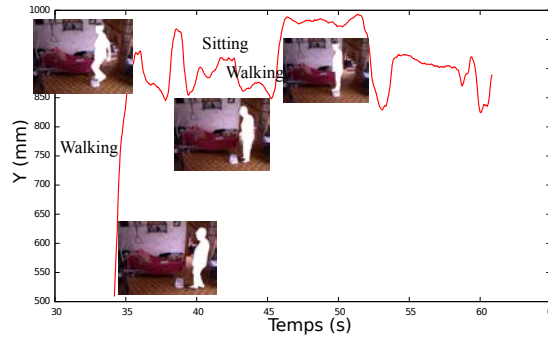


FIGURE 9.3 – Détection de l'activité « Marche », pour une personne de 81 ans marchant à son domicile.

l'algorithme de l'état dans lequel se trouve la personne. Les situations où la personne s'assoit et où elle se penche sont correctement reconnues. Nous pouvons également constater que les phases où la personne marche avant et après la réalisation de l'activité « Assis » et « Penché » sont reconnues comme « Marche ». L'activité « Monté » n'est pas reconnue. La personne montait sur un marche pied de 13,5 cm. La personne a tout d'abord essayé de monter mais est retombée au sol très rapidement, n'arrivant pas à tenir sur le marche pied. Puis, elle est remontée et est restée immobile quelques secondes. La montée n'était peut être pas assez grande pour que l'algorithme la détecte. Dans la base de donnée, pour la situation « Monté », les personnes montaient sur une chaise et seulement 5 d'entre elles ont réalisé une deuxième séquence avec un marche pied. Nous pensons que les paramètres appris pour une chaise sont différents de ceux qui seraient appris pour un marche pied. C'est pourquoi la situation n'a pas été reconnue.



(a) Trajectoire du centre de masse pour une personne s'asseyant. (b) Trajectoire du centre de masse pour une personne se penchant.



(c) Trajectoire du centre de masse pour une personne montant sur un marche pied.

FIGURE 9.4 – Représentation du centre de masse, sur l'axe vertical, pour différentes activités réalisées par une personne âgée de 81 ans.

9.3 Conclusion

Les différents tests effectués avec des personnes âgées permettent de nous donner des premiers résultats de ce que nous pourrions obtenir si nous réalisions des expérimentations avec elles. Ainsi, nous pouvons constater que les trajectoires du centre de masse, sur l'axe vertical, sont plus irrégulières que ce que nous obtenions avec des personnes plus jeunes. Les courbes suivent, tout de même, une trajectoire sinusoïdale, ce qui nous permet d'extraire les paramètres de la marche. Toutefois, nous ne pouvons pas en conclure que les paramètres extraits sont précis,

n'ayant pas de système de référence lors des tests. Concernant la détection d'activité, les tests nous montrent que nous devons modifier la base de données sur laquelle les modèles sont appris, avant de nous lancer dans une expérimentation, notamment pour la situation « Monter sur un obstacle ».

Conclusion

Différentes expérimentations ont été mises en place pour tester les algorithmes d'analyse des paramètres de la marche et de détection des activités. Les expérimentations concernant la détection d'activité nous ont permis d'identifier les modèles, que nous avons construits, les plus proches de la réalité. Nous avons également constaté que certains d'entre eux pouvaient s'adapter à des situations n'appartenant pas à la base d'apprentissage. Notamment, certains modèles se sont adaptés aux situations d'occlusions, qui sont des situations très souvent rencontrées dans une maison. Les expérimentations testant la précision des paramètres de la marche fournis par l'algorithme nous ont révélé que des erreurs subsistent, plus ou moins grandes selon les situations réalisées. La caméra Kinect étant précise à quelques centimètres près, comme nous l'avons précisé à la section 3.3, nous nous approchons de cette erreur.

L'expérimentation réalisée avec les personnes âgées nous donne une idée du type de trajectoires que nous pouvons obtenir. Mais les tests effectués sont insuffisants pour pouvoir conclure que le système est adapté aux personnes âgées. De plus, avant de tester sur les personnes âgées, nous devons compléter la base d'apprentissage. Actuellement, les paramètres des modèles sont appris à partir d'une base de données contenant uniquement des images de sujets jeunes.

Le système sera, dans les prochains mois, testé dans des conditions réelles et avec des personnes âgées dans le cadre du projet SATELOR. Ce projet, financé par la région Lorraine, regroupe des industriels, des hospitaliers et des chercheurs. L'objectif est d'industrialiser un système permettant de maintenir la personne à son domicile. La prochaine étape de ce projet est d'installer le système, dans un premier temps, dans une chambre de l'OHS (Office Hygiène Sociale) à Nancy et d'enregistrer l'activité d'un patient pour pouvoir, par la suite, traiter les données avec notre système. Nous pourrons ainsi constater les manques et atouts du système en situation réelle. Le but final étant d'installer le système au domicile des personnes âgées pour apprendre leurs habitudes, évaluer l'évolution de leur état au quotidien et détecter les éventuelles chutes.

Quatrième partie

Système pour le maintien à domicile

Introduction

La reconnaissance d'activité et l'extraction des paramètres de la marche sont les deux axes de recherche que nous avons développés pour traiter le problème de la sécurisation de la personne à son domicile et de l'évaluation de son degré d'autonomie. Dans les parties précédentes, nous avons présenté la méthode et les résultats obtenus pour chacun des deux problèmes de manière séparée, tout en sachant que les deux contributions s'appuient sur une base commune de caractéristiques extraites des images de profondeur. L'objectif final est de développer un système unique pour le maintien à domicile.

Nous présentons, dans le premier chapitre de cette partie, notre vision du système final, en intégrant la détection d'activité et l'analyse de la marche dans un seul système. Par exemple, lorsque l'algorithme détecte que la personne est dans l'état « Marche », l'algorithme analysant les paramètres de la marche est activé. Nous proposons également une méthode pour apprendre les habitudes de la personne et ainsi, détecter toutes anomalies suite à un événement inhabituel. La méthode consiste à cartographier l'environnement des habitudes de la personne.

Dans un deuxième chapitre, nous indiquons d'autres problématiques pour lesquelles le système pourrait être adapté. Par exemple, ce système peut servir à discriminer plusieurs personnes dans la scène et ainsi, suivre l'évolution de l'état de fragilité de plusieurs personnes en même temps. Chaque personne pourrait être identifiée à travers un certain nombre de caractéristiques, comme sa taille, la forme de sa trajectoire de son centre de masse, etc. Ce système pourrait également prévenir, en plus des chutes et de la perte de fragilité des personnes, les troubles cognitifs. En effet, certaines études ont montré qu'une vitesse lente pouvait être un indicateur de troubles cognitifs et de démences à venir.

Chapitre 10

Vers une implantation au domicile

Sommaire

10.1 Couplage des fonctionnalités actuelles du système	127
10.2 Cartographie des habitudes dans l'environnement	129
10.2.1 Limitations du système et contraintes de l'environnement	129
10.2.2 Détection des cas d'occlusions partielles de la personne	130
10.2.3 Cartographie des habitudes	133
10.2.4 Détection des anomalies	134
10.3 Conclusion	135

Actuellement, le système développé est conçu pour reconnaître différentes activités et extraire des paramètres de la marche. Dans ce chapitre, nous présentons la manière dont nous avons imaginé intégrer les deux fonctionnalités pour former un système unique, répondant à la problématique du maintien à domicile des personnes âgées. Nous présentons, dans un premier temps, comment coupler les fonctionnalités actuelles du système pour sécuriser et évaluer le degré d'autonomie de la personne. Puis, nous proposons une méthode pour apprendre les habitudes des personnes.

10.1 Couplage des fonctionnalités actuelles du système

Nous avons précisé, dans l'introduction de cette thèse, que pour sécuriser la personne, nous avons besoin d'utiliser un algorithme de détection de chutes et d'activités à risque. Concernant l'évaluation du degré d'autonomie de la personne, nous avons précisé que les activités réalisées par la personne et les évolutions dans les paramètres de la marche seraient prises en compte. L'idée est de coupler les algorithmes de détection d'activités et d'extraction des paramètres de la marche pour faire un système unique répondant à la problématique du maintien à domicile des personnes âgées.

Nous avons imaginé mettre au domicile une ou plusieurs caméras. L'algorithme détectant l'activité de la personne permettrait, en temps réel, de savoir quelle activité la personne réalise. Selon l'activité qui sera reconnue, l'algorithme effectuera telle ou telle tâche, comme montré à la figure 10.1. Les différentes tâches sont dépendantes de l'activité identifiée :

- « Marche » : enregistrement du temps passé dans cette situation, que l'on nommera activité « active », et extraction des paramètres de la marche quand la personne est entièrement visible ;

- « Debout » : enregistrement du temps passé dans cette situation nommée activité « active » ;
- « Chute » et « Allongé Sol » : envoi d'une alerte à la famille ou aux secours ;
- « Monté » et « Accroupi » : envoi d'un signal à la famille ou au médecin pour avertir que la personne adopte des comportements dangereux pouvant la conduire à des situations telles que la chute ;
- « Assis » et « Allongé Canapé » : enregistrement du temps passé dans cette situation que l'on nommera activité « passive ».

Ainsi à la fin de chaque journée, le système pourra fournir un résumé des informations concernant :

- le cumul du temps passé dans une activité « passive » ;
- le cumul du temps passé dans une activité « active » ;
- le nombre de changements d'activité ;
- la moyenne des longueurs de pas, cadence et vitesse de marche de la journée.

Il sera alors possible d'observer l'évolution dans le temps (dans la semaine, le mois ou l'année) de ces indicateurs. Cette information, couplée à l'évolution des paramètres de la marche de la personne, est un critère d'évaluation de la perte ou non d'autonomie de la personne. L'alerte automatique aux proches ou aux professionnels de la santé, en cas de chutes et d'activités à risque, permettra de sécuriser la personne au domicile. Les différents algorithmes construits permettent donc d'apporter un ensemble d'informations, qui couplées ensemble, pourront aider à maintenir la personne plus longtemps à son domicile.

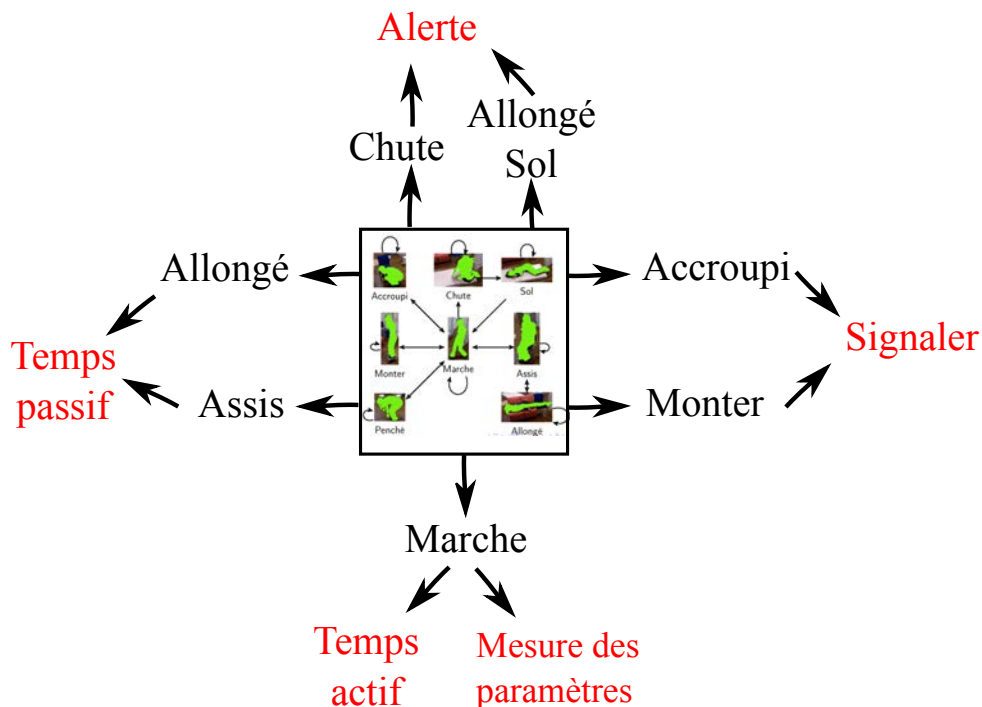


FIGURE 10.1 – Les différentes actions réalisées par l'algorithme selon l'état dans lequel se trouve la personne d'après le MMC.

10.2 Cartographie des habitudes dans l'environnement

L'objectif dans cette section est de développer une méthode d'apprentissage des habitudes des personnes dans l'environnement. Plus précisément, nous voulons cartographier les activités courantes de la personne selon l'endroit de la pièce dans lequel elle se trouve. Ainsi, l'évolution des habitudes au cours des semaines ou des mois peut constituer un indicateur de modification de l'état de la personne, par exemple si elle reste plus longtemps allongée qu'à son habitude.

Pour constituer la carte des habitudes de la personne, nous devons relever ses activités au cours du temps. Lors de l'apprentissage de cette carte, nous devons tenir compte du fait que quand la personne n'est pas entièrement visible, l'activité qu'elle réalise n'est pas certaine. Par exemple, quand la personne se trouve derrière une table à moitié visible nous ne pouvons pas connaître avec certitude son activité. Nous devons différencier dans la carte les zones où la personne est vue entièrement des zones où elle peut être cachée.

10.2.1 Limitations du système et contraintes de l'environnement

Les occlusions peuvent être de différentes natures. Comme nous l'avons étudié à la section 8.2, les occlusions peuvent être dues aux meubles de la pièce. En effet, les tables ou les chaises peuvent, par exemple, bloquer la visibilité de la personne. De la même façon, en bordure du champ de vision de la caméra, la personne n'est que partiellement visible. Donc, les problèmes d'occlusions, dues aux meubles et aux limites du champ de vision de la caméra, sont à prendre en compte lorsque l'on suit la trajectoire d'une personne. En effet, si l'analyse des activités est effectuée lorsque la personne est à moitié cachée l'activité ne sera peut être pas reconnue, comme nous l'avons vu à la section 8.2. La figure 10.2 illustre le cas d'une personne marchant à la limite du champ de vision de la caméra. L'image couleur nous montre que la personne est à moitié détectée. Son côté gauche se trouve en dehors du champ de vision de la caméra.

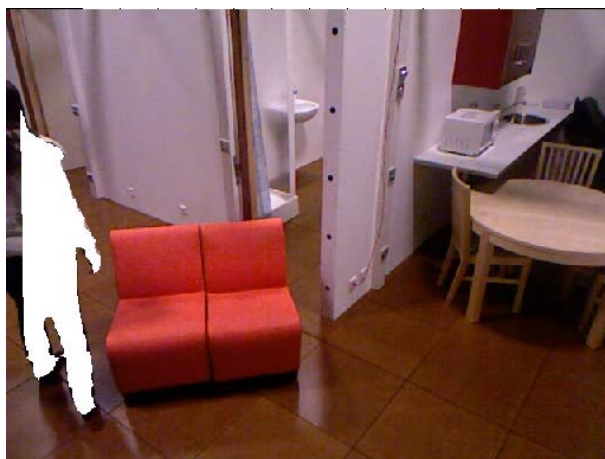


FIGURE 10.2 – Personne sortant du champ de vision et étant à moitié détectée.

La section 8.2 permet de constater que certains de nos modèles de Markov cachés sont robustes aux occlusions, en classifiant correctement les activités, dans des conditions non rencontrées lors de l'apprentissage. La plupart des modèles ont le même problème qui est de confondre les trois activités « Penché », « Accroupi » et « Chute ». Ainsi, même s'il est possible de reconnaître certaines activités dans les zones d'occlusion il est important d'identifier ces zones dans lesquelles les performances du système sont différentes. Les zones de sortie du champ de

vision sont également à prendre en compte dans l'interprétation des résultats produits par le système. Ces zones de sorties se situent en bord d'image mais également dans les franchissements de portes si celles-ci se trouvent dans le champ de vision.

Nous devons donc différencier les moments où la personne est cachée, des situations où elle est entièrement visible pour intégrer cette information dans la carte des habitudes.

10.2.2 Détection des cas d'occlusions partielles de la personne

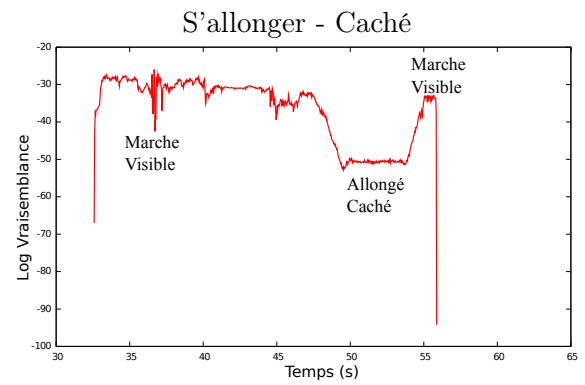
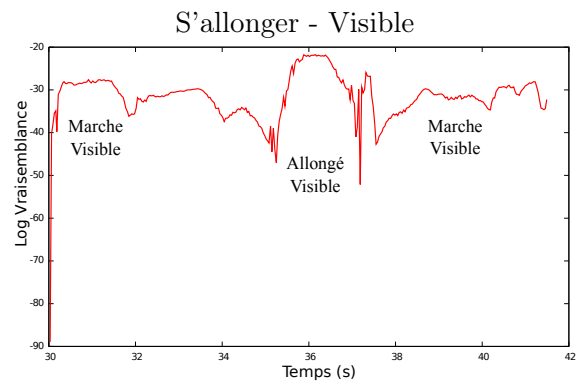
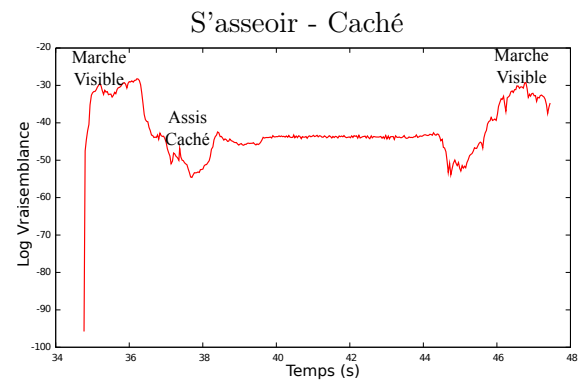
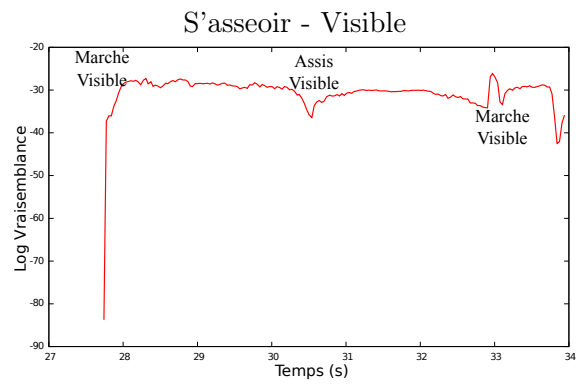
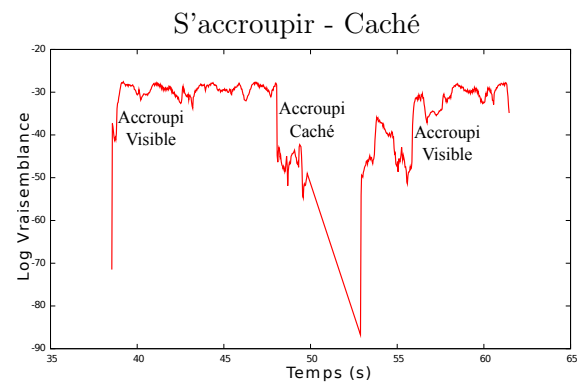
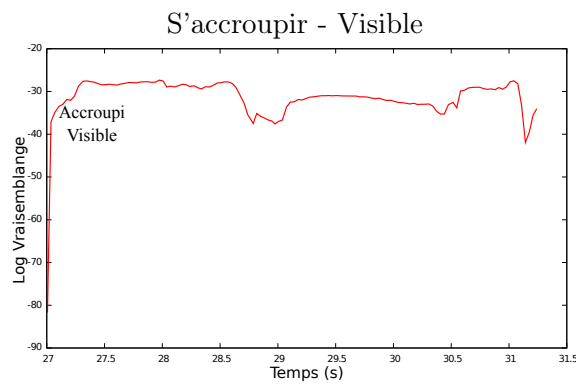
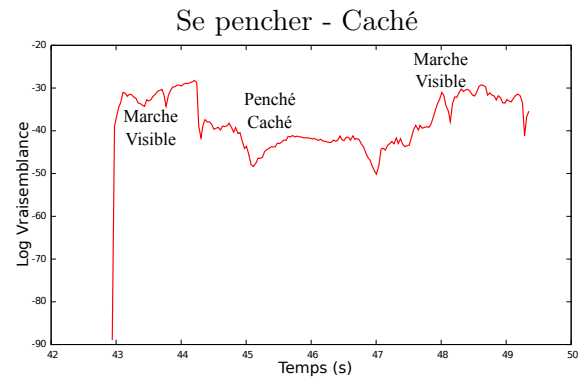
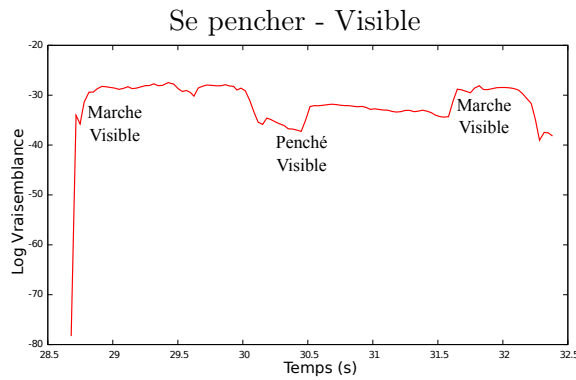
La méthode proposée pour détecter les cas d'occlusions est fondée sur le calcul de la vraisemblance permettant de savoir si le modèle correspond aux observations. La vraisemblance est une mesure de la probabilité des observations O sachant le modèle λ , soit $P(O|\lambda)$. Elle est calculée à partir des probabilités obtenues dans la phase *Forward* de l'algorithme *Forward-Backward*. Pour rappel, la phase *Forward* permet de calculer le α pour chaque état, à chaque instant. La vraisemblance correspond au calcul suivant :

$$Vraisemblance(t) = \log \sum_{i=1}^N \alpha_t(i).$$

La vraisemblance à l'instant t est égale à la somme des probabilités α à t pour tous les états. Le logarithme de la vraisemblance est préféré pour représenter les résultats.

L'idée ici est de comparer les valeurs prises par la vraisemblance lorsque la personne est entièrement visible, avec les valeurs prises lorsque la personne n'est pas entièrement visible. Nous avons alors demandé à une personne d'effectuer les huit activités (les mêmes qui étaient utilisées pour construire les MMC) pour obtenir l'échantillon des valeurs prises par la vraisemblance lorsque la personne effectue les différentes activités en étant entièrement visible. Ensuite nous avons demandé à la personne de réaliser les séquences derrière des obstacles pour obtenir un échantillon de valeurs prises par la vraisemblance lorsque la personne est cachée. Chaque séquence a été analysée avec le MMC à 8 états et 3 observations. Ces situations correspondent aux situations décrites au chapitre 8. Les valeurs de la vraisemblance pour chaque activité, réalisée en condition « Caché » et entièrement « Visible », sont représentées à la figure 10.3. Chaque courbe correspond aux valeurs de la vraisemblance à chaque pas de temps. Nous avons inscrit sur chaque graphe les situations réalisées par le sujet. Par exemple, pour le graphe correspondant à la situation « Marcher - Caché », la personne marche en étant visible jusqu'à 44,1 s, puis continue de marcher derrière des chaises en étant à moitié visible.

En comparant les situations réalisées lorsque la personne est visible et cachée, nous pouvons voir que les valeurs prises par la vraisemblance sont plus petites quand la personne est cachée. En d'autres termes, le modèle correspond mieux aux observations quand la personne est entièrement visible. Ce constat est logique puisque le modèle a appris sur une base de données ne contenant que des situations où la personne est entièrement visible. Nous pouvons également constater que les valeurs de la vraisemblance effectuent un pic vers le bas lorsque la personne passe d'une activité où elle est entièrement visible à une activité où elle est en partie cachée. Par exemple, pour la situation « Marcher - Caché », le passage de l'état visible à caché est très net par rapport aux valeurs de la vraisemblance. Dans le cas des situations réalisées en étant visibles, nous pouvons également constater une descente des valeurs de la vraisemblance, mais moins prononcée, lorsque la personne change d'activité. Par exemple, les situations « S'accroupir - Visible » ou « S'asseoir - Visible » ont une légère descente des valeurs de la vraisemblance lorsque la personne passe de l'activité « Marche » à « Accroupi » ou « Assis ». Pour la situation « S'accroupir - Caché » et « Chuter+Allongé au sol - Caché », nous voyons que la courbe de la vraisemblance ne cesse de diminuer à un moment. Ce passage correspond au moment où la personne est entièrement cachée et où le modèle ne reçoit plus aucune information.



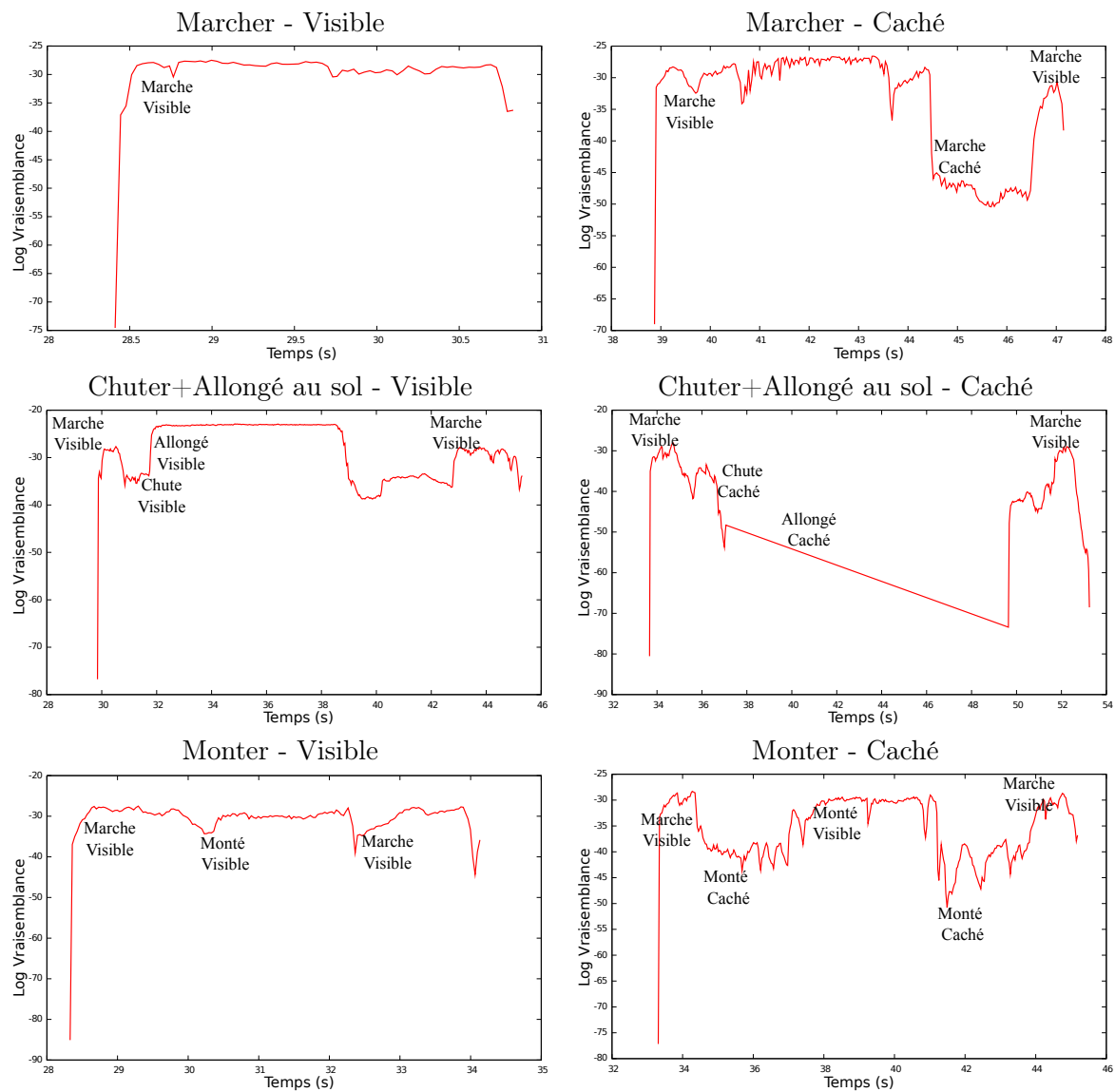


FIGURE 10.3 – Courbes des vraisemblances pour différentes activités réalisées par un sujet entièrement et partiellement visible.

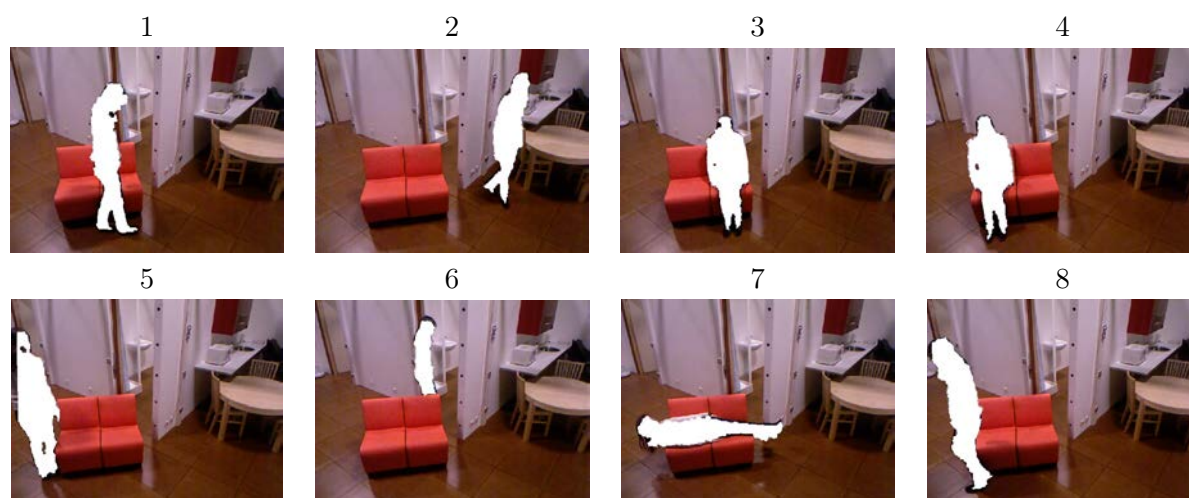


FIGURE 10.4 – Séquence dans laquelle le sujet réalise plusieurs activités.

Donc, à partir des valeurs de vraisemblance, il est possible de reconnaître une situation où la personne est cachée en utilisant un seuil en dessous duquel la personne sera considérée cachée par un obstacle. Ainsi, nous pouvons discriminer les situations où la personne est entièrement visible et celles où elle ne l'est pas.

10.2.3 Cartographie des habitudes

Pour apprendre l'environnement, nous découpons la scène en zones de 50 cm de côté. Nous demandons à une personne d'effectuer plusieurs activités dans l'appartement expérimental. Nous y avons placé deux fauteuils mis côte à côte. Ainsi, la personne a pu s'asseoir, s'allonger, marcher derrière les fauteuils en étant cachée et marcher devant les fauteuils en étant visible. Cette séquence est montrée à la figure 10.4. Les images représentent les différentes activités réalisées par la personne. Le corps de la personne est toujours en partie visible durant toutes ces activités.

L'algorithme apprend, pour chaque zone de l'environnement, la fréquence des activités réalisées par la personne, la visibilité ou non du corps de la personne et les bords du champ de vision de la caméra. Plus précisément, l'algorithme observe si la personne est cachée ou entièrement visible, par rapport à la valeur prise par la vraisemblance. Si la personne est visible, le modèle apprend pour la zone de l'environnement où elle se trouve qu'il est possible, à cet endroit, d'effectuer l'activité détectée. Si la personne est cachée par un meuble, le modèle apprend que la case, où elle se trouve, est localisée derrière un obstacle. Pour finir, nous considérons un bord comme étant l'endroit où la personne sort entièrement du champ de vision de la caméra.

La figure 10.5 représente l'apprentissage de l'environnement qui a été construit lorsque la personne a effectué les différentes activités dans l'appartement expérimental. Les cases blanches correspondent aux zones non explorées. La personne n'a jamais été détectée à ces endroits. La couleur bleue représente les zones où la personne a été détectée comme réalisant une activité dite « passive ». Nous avons représenté une chaise dans cette case, pour montrer que dans cette zone se trouve une chaise, car le MMC a détecté que la personne était soit dans l'état « Allongé Canapé », soit dans l'état « Assis ». La couleur rouge représente les zones où la personne a été détectée comme étant cachée par un obstacle. La couleur verte représente les zones où la personne a été détectée réalisant une activité dite « active ». Le MMC a détecté que la personne était soit dans l'état « Marche », soit dans l'état « Debout ». Les proportions des couleurs dans chaque

case représentent les proportions des situations détectées. Nous pouvons voir que les zones rouges se situent correctement derrière les deux fauteuils. Les zones bleues, représentant les activités « passives », sont situées à l'endroit où se trouvent les deux fauteuils. Cela nous montre que les activités « Allongé Canapé » et « Assis » ont été correctement identifiées au niveau des deux fauteuils. Les zones vertes, représentant les activités « actives », sont situées à l'endroit où la personne a marché en étant entièrement visible. Sur la carte, nous pouvons également voir l'inscription « bord », placée là où la personne est sortie entièrement du champ de vision de la caméra.

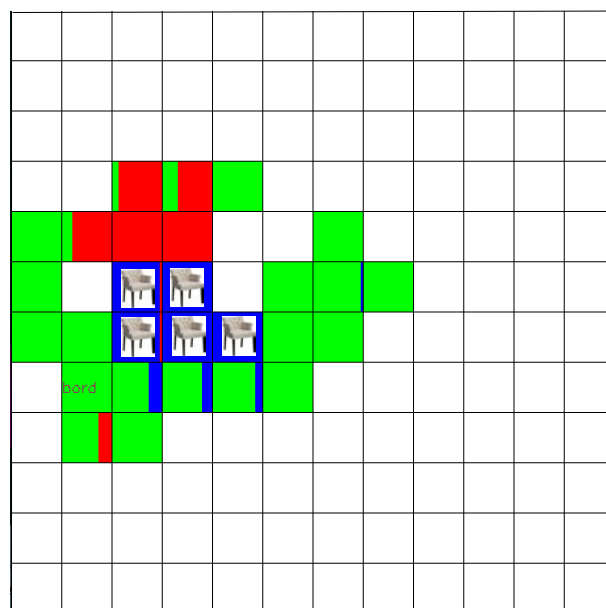


FIGURE 10.5 – Carte de l'environnement représentant les habitudes apprises d'une personne (en blanc : zones non explorées, en bleu : zones passives, en vert : zones actives et en rouge : zones se trouvant derrière un meuble).

Dans la réalité, au domicile des personnes âgées, nous imaginons apprendre au fur et à mesure les activités réalisées par la personne pour construire la carte de leur environnement et de leurs habitudes.

10.2.4 Détection des anomalies

La carte peut être utilisée comme une méthode de détection d'anomalies. La cartographie des habitudes de la personne permet d'identifier, à différents intervalles de temps par exemple, l'évolution des activités de la personne. L'évolution des habitudes peut être un indicateur d'une éventuelle anomalie dans l'état général de la personne.

D'autres anomalies peuvent être détectées, en utilisant la carte, pouvant faire l'objet d'une alerte envoyée aux personnes extérieures. Notamment, lorsque la personne disparaît du champ de vision, la carte peut renseigner sur sa proximité avec une zone d'occlusion et nous indiquer un risque de chute non visible. Nous détectons les moments où la personne est entièrement cachée après avoir visité une case rouge. L'algorithme calcule le temps passé dans cette situation. Si le temps est supérieur à un certain seuil (de quelques secondes), nous en déduisons une possibilité de chute. La figure 10.6 correspond à une situation où une personne se trouve cachée derrière un fauteuil. Sur l'image de l'apprentissage de l'environnement en haut à gauche, le cercle représente

la dernière position du centre de masse qui a été détectée lorsque la personne était encore visible. Nous pouvons voir que la position du cercle se trouve au dessus d'une case rouge, signifiant que cette zone a été apprise comme étant une case où la personne peut potentiellement être cachée. Un compteur est lancé à partir du moment où la personne n'a plus été visible, pour déterminer s'il y a chute ou non.

A l'inverse, la carte peut nous aider à prendre en compte les zones de sorties habituelles du champ de vision. La personne peut sortir du champ de vision en bord d'image mais aussi suite au franchissement d'une porte qui serait visible dans la scène. Si la personne disparaît du champ de vision à proximité d'une zone de sortie habituelle, l'alerte ne sera pas déclenchée.

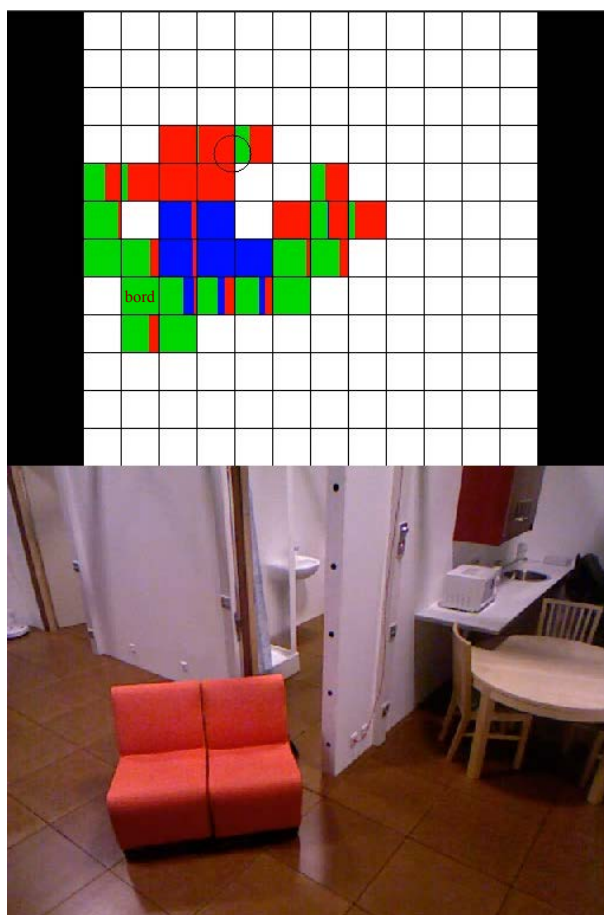


FIGURE 10.6 – Carte de l'environnement sur laquelle est représentée la dernière position du centre de masse perçue par le système.

10.3 Conclusion

Le couplage entre la fonctionnalité de détection des activités et la fonctionnalité permettant d'extraire les indicateurs de marche permet d'obtenir de nombreuses données qui peuvent servir à sécuriser la personne à son domicile et évaluer son degré de fragilité. En mettant des outils en place, tels que la détection des occlusions et l'apprentissage de l'environnement, nous renforçons la robustesse du système pour permettre son fonctionnement au domicile, dans un environnement où les occlusions de la personne sont très courantes.

La carte de l'environnement permet de connaître les zones d'occlusions habituelles en cumulant au cours du temps l'analyse de la vraisemblance. La détection d'une occlusion peut ensuite se faire en regardant la position de la personne sur la carte, en étant robuste aux éventuelles variations ponctuelles de vraisemblance. Elle peut être utilisée pour éviter d'éventuelles fausses détections, en prenant en compte les zones d'occlusions et les activités habituelles dans la zone où l'alerte a été détectée. Enfin, la carte de l'environnement donne une représentation spatiale des habitudes qui peut être utilisée pour analyser plus finement les évolutions des habitudes de la personne.

Chapitre 11

Autres perspectives d'utilisation du système

Sommaire

11.1 Différenciation des personnes	137
11.2 Couplage de plusieurs capteurs	138
11.3 Observance des patients	139
11.4 Prévention de certains troubles	139
11.5 Conclusion	140

Le système est actuellement conçu pour apporter une solution à la problématique du maintien à domicile des personnes âgées. Nous pensons que ce système pourrait avoir d'autres applications. Nous proposons ici quelques perspectives d'évolution et exemples d'utilisation du système.

11.1 Différenciation des personnes

Le système actuel est prévu pour qu'il n'y ait qu'une seule personne dans le champ de vision de la caméra. Lorsque plusieurs personnes entrent dans le champ de vision de la caméra, le système analyse la marche et l'activité de l'une des personnes, qui n'est pas nécessairement la personne que l'on souhaite suivre. Ces erreurs d'attributions provoqueront des données aberrantes, pouvant gêner l'analyse de l'évolution des paramètres de la marche de la personne à suivre. Nous aimerions donc pouvoir distinguer les personnes entre elles pour analyser l'activité et la marche de la ou les personnes à suivre.

Nous nous basons sur l'idée que chaque personne a une façon de marcher différente. Nous pourrions alors identifier chaque personne en prenant en compte un certain nombre de caractéristiques telles que sa taille, sa longueur de pas moyenne, sa cadence, sa vitesse de marche, la symétrie de sa marche, etc. La taille de la personne est une information que nous connaissons, car nous avons une méthode d'extraction du sol, comme présenté à la section 4.1, et une méthode de détection du point maximum du corps. Les longueurs, cadences et vitesses de marche sont également connues avec l'algorithme d'analyse de la marche. Et la symétrie des pas peut être obtenue en analysant la régularité des trajectoires du centre de masse. Nous pensons qu'il serait alors possible, à l'installation du système, d'apprendre les caractéristiques de la personne et de les mettre à jour au cours des évolutions. Ainsi, l'algorithme pourrait suivre plusieurs personnes en leurs attribuant les paramètres observés.

11.2 Couplage de plusieurs capteurs

Nous imaginons poser plusieurs caméras aux domiciles des personnes. Le positionnement des caméras serait effectué pour qu'il n'y ait pas de zone de recouvrement, car nous avons montré, au chapitre 7, que l'interférence entre caméras dégrade la précision des résultats. Par exemple, dans l'appartement expérimental, nous pensons installer les caméras comme montré à la figure 11.1. Une caméra serait installée pour couvrir le salon ayant une profondeur d'environ 7 m, une autre pour la cuisine et une dernière pour la chambre. Les caméras n'ayant pas de zone de recouvrement, certains endroits ne seront pas couverts. Les zones importantes où sont situés la table, les chaises, les fauteuils, le canapé et le lit doivent être vues. Une solution pour pallier au problème des zones non vues est d'effectuer les traitements concernant l'analyse de l'activité et l'apprentissage de l'environnement pour chaque caméra. Ainsi, lorsqu'une personne ne sera visible d'aucune caméra, sa dernière position visible sera recherchée. La carte de l'environnement de la caméra, où la personne est apparue visible, nous permettra de connaître la distance entre la personne et une zone d'occlusion ou une zone de bord du champ de vision. Si la personne était proche d'un bord, nous pourrions en conclure qu'elle se trouve, sûrement, entre deux champs de vision de caméra. Comme précédemment, nous pourrions ajouter la notion de temps et apprendre les habitudes de la personne. C'est-à-dire que nous pourrions apprendre combien de temps, la personne passe, généralement, entre chaque zone non visible par la caméra, pour savoir si la situation est normale.

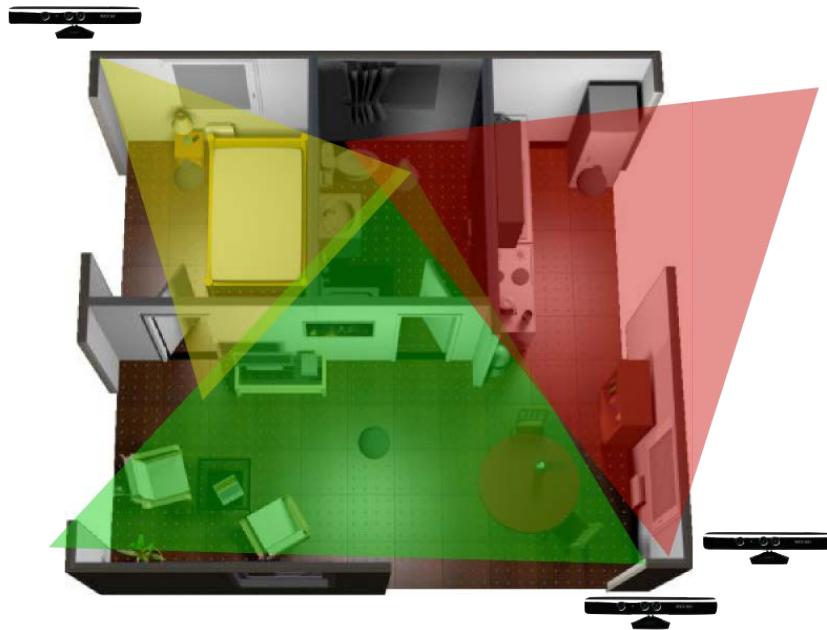


FIGURE 11.1 – Proposition d'emplacement de plusieurs caméras permettant de couvrir tout un appartement.

Nous n'avons pas mis de caméra dans la salle de bains et les toilettes, à la figure 11.1, pour préserver l'intimité de la personne. L'idée serait de mettre d'autres capteurs, tels que des détecteurs d'ouvertures sur les robinets pour savoir si la personne est active. Ainsi, si aucune activité n'est détectée dans la pièce, nous pourrions en déduire un risque d'activité anormale.

11.3 Observance des patients

Nous pensons que ce système pourrait servir au médecin, pour vérifier que le patient respecte, par exemple, les indications qui lui ont été données sur sa façon de marcher. En d'autres termes, le médecin pourrait évaluer l'observance du patient. Nous avons testé le système avec une personne âgée de 90 ans, ayant des problèmes d'équilibre. Nous lui avons demandé de marcher sans sa canne et avec sa canne. Quand la personne marche sans sa canne, la trajectoire du centre de masse n'est pas régulière comme illustré à la figure 11.2(a). A l'inverse, quand elle marche avec sa canne, la courbe du centre de masse suit une trajectoire sinusoïdale plus régulière, comme le montre la figure 11.2(b). Le médecin de cette personne lui a conseillé de marcher uniquement avec sa canne au domicile pour lui éviter de chuter et pour ne pas empirer ses problèmes d'équilibre. Ce système lui permettrait donc de pouvoir observer, à travers la régularité, ou non, de la trajectoire du centre de masse, si la personne respecte ses conseils. Ainsi, il pourrait prévenir les futures évolutions de l'état de la personne liées à la non utilisation de la canne, alors qu'actuellement le médecin constate les dégradations lorsque la personne chute.

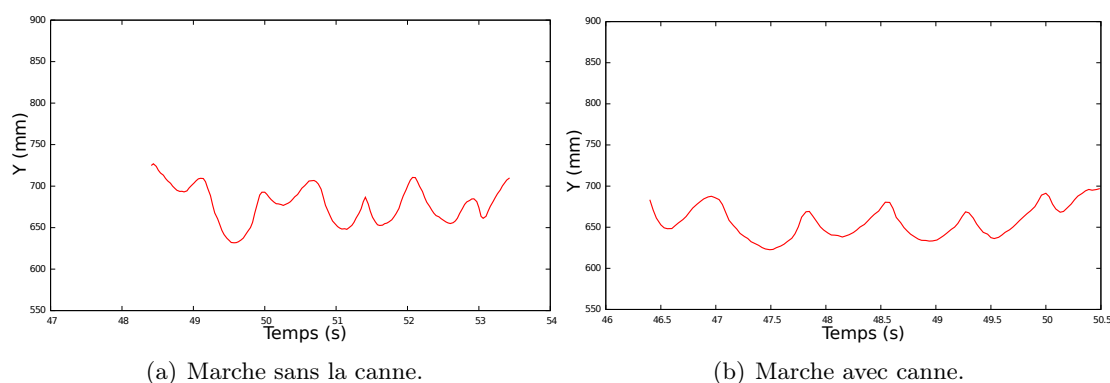


FIGURE 11.2 – Représentation du centre de masse, sur l'axe vertical, d'une personne de 90 ans ayant des problèmes d'équilibre .

En restant dans le même contexte, cet outil pourrait également être utile pour que le médecin évalue l'observance dans le cas des patients en rééducation, suite à un accident ou une amputation par exemple. Un des problèmes en rééducation est le retour des patients au domicile. Les personnes à l'hôpital suivent les exercices qui leur sont indiqués par le personnel soignant. Mais, une fois de retour chez eux, ils ne continuent pas l'entraînement instauré par le médecin. L'installation d'un tel système au domicile permettrait au médecin de revoir avec le patient chaque semaine, chaque mois, les trajectoires de son centre de masse. Il pourrait alors savoir si son patient continue la rééducation en marchant correctement chez elle et continue de progresser.

11.4 Prévention de certains troubles

L'analyse de la marche est utilisée, dans le système actuel, comme indicateur de fragilité et de chute à venir. A la section 5.3, nous avons présenté certaines études montrant que l'analyse des troubles de la marche peut aider à détecter les risques d'hospitalisations, les troubles cognitifs et les démences. Nous pouvons imaginer, tout comme pour la prévention de la fragilité, d'évaluer au quotidien les paramètres de la marche et d'observer leur évolution dans le temps. Les personnes seraient alors orientées, assez rapidement, vers les spécialistes adéquates. Comme il a été dit à

la section 5.3.3, dans le cas des démences, l'un des problèmes majeurs est que bien souvent les personnes sont prises en charge trop tard. Avec un système analysant la marche au quotidien, nous pouvons alerter dès les premiers signes de troubles et les orienter vers des neuropsychologues par exemple, pour leur faire passer des tests cognitifs.

Pour rendre le système capable de prévenir ces différents troubles, nous devons mettre en place des expérimentations pour connaître les seuils à partir desquelles nous alerterons les médecins. Aujourd'hui, il y a peu d'information à ce sujet, car l'évaluation de la marche n'est pas une pratique courante pour détecter un problème cognitif ou un risque d'hospitalisation.

11.5 Conclusion

Ce système a été pensé pour répondre au besoin du maintien à domicile des personnes âgées. La détection de chutes des personnes âgées au domicile est une thématique pouvant concerner d'autres populations comme les personnes handicapées ou même les personnes épileptiques, dont un des symptômes possibles est la perte de connaissance entraînant des chutes. La chute est souvent brutale et peut entraîner des conséquences physiques empêchant la personne de se relever. Nous pouvons également imaginer que l'analyse de la marche puisse être détournée de sa fonction première qui était de prévenir les chutes et la fragilité des personnes âgées. L'analyse de la marche, au quotidien au domicile, peut être un outil supplémentaire pour le médecin pour prévenir les risques d'hospitalisations et de démences. Cet outil peut également permettre d'analyser les anomalies de symétrie de la marche et ainsi fournir une indication concernant l'évolution des progrès d'une personne suite à un accident par exemple.

Sur le plan technique, nous avons identifié la possibilité d'utiliser les mêmes algorithmes pour distinguer et suivre plusieurs personnes en constatant que les caractéristiques extraites sont spécifiques à chacun. Une autre perspective d'évolution est le couplage de plusieurs caméras pour permettre une meilleure couverture du domicile. Ce couplage pose la question du non recouvrement des champs de vision et de la gestion des zones aveugles entre les caméras. Enfin, la fusion de l'information provenant de capteurs complémentaires installés dans des pièces telles que la salle de bains ou les toilettes permettrait de compléter la zone de surveillance tout en conservant l'acceptabilité du système.

Conclusion

Cette partie permet de montrer le couplage des deux contributions de la thèse, la détection d'activités et l'analyse de la marche, pour former un système permettant le maintien à domicile des personnes âgées. Le système pourrait envoyer une alerte aux personnes extérieures en cas de détection de l'activité « Chute » et des activités à risque, « Accroupi » et « Monté ». Il pourrait, également, évaluer au quotidien les paramètres de la marche de la personne et comparer ses activités à celles réalisées habituellement, dans le but d'évaluer l'état de fragilité de la personne. Ce système peut fonctionner pour une personne entièrement visible et cachée par des meubles. Une carte de l'environnement, représentant les habitudes des personnes et découpant la scène en zones d'activités possibles à cet endroit, est ajoutée au système pour le rendre plus robuste dans les cas des zones partiellement occultées par des meubles.

Ce système peut également avoir d'autres applications. Il pourrait nous servir à distinguer les personnes. Actuellement, une seule personne est détectée dans la scène. Dans le cas où plusieurs personnes sont dans la scène, le système pourrait suivre la personne qui nous intéresse parmi d'autres, ou en suivre plusieurs et distinguer les informations appartenant à chacune d'entre elles. Cette distinction se ferait par l'apprentissage de caractéristiques discriminant chaque personne. Une autre utilité, que nous avons évoquée dans ce chapitre, est la possibilité de coupler les caméras entre elles et avec d'autres capteurs pour ainsi avoir des informations dans toutes les pièces de la maison. Nous avons, également, évoqué dans cette partie la possibilité d'utiliser ce système comme outil de suivi de l'observance pour les médecins. Ce système pourrait, aussi, prévenir d'autres risques comme les troubles cognitifs ou les risques d'hospitalisation.

Conclusion générale et perspectives

L'augmentation du nombre de personnes âgées est un enjeu majeur de santé publique des prochaines années. Le maintien à domicile s'avère être une préoccupation importante pour les personnes âgées et leur famille. Mais le domicile n'est pas un endroit sans risque pour les personnes, il peut même être responsable de leur perte d'autonomie. Une des causes fréquentes de perte d'autonomie au domicile est la chute. Les séquelles physiques peuvent conduire la personne au handicap ou à des incapacités locomotrices, entraînant une perte d'autonomie pouvant mener au placement en institution. Les séquelles psychologiques sont aussi très fréquentes après une chute. Le risque de séquelles psychologiques augmente avec le temps passé au sol. Elles provoquent la diminution des activités et des sorties par peur de rechuter. Ces séquelles ont pour effet d'entraîner les personnes vers une perte d'autonomie et de compromettre leur maintien à domicile. Pour que les personnes âgées puissent rester le plus longtemps possible autonomes à leur domicile, il est nécessaire de les sécuriser chez elles en détectant les chutes pour intervenir le plus rapidement possible et diminuer les risques de séquelles physiques et psychologiques. En plus de détecter les chutes, nous pensons qu'il est également nécessaire de prévenir les chutes et plus généralement de prévenir les risques de fragilité de la personne. Actuellement, à l'hôpital, il est possible de mesurer, lors des consultations, le niveau de dépendance des personnes à l'aide de tests cliniques, analysant la marche de la personne. Ces mesures sont alors faites dans des conditions modifiées comparées aux mesures qui pourraient être faites au domicile. Notre objectif est donc de proposer un système permettant de détecter les chutes pour éviter l'aggravation des séquelles physiques et psychologiques et détecter, dans le milieu écologique, une détérioration du comportement locomoteur pour ainsi prévenir la fragilité des personnes. Ainsi, nous aurons une mesure quotidienne pour évaluer l'évolution de la personne.

Dans cette thèse, nous proposons un système, pour détecter les chutes et prévenir la fragilité des personnes âgées. Pour détecter les chutes, nous avons développé un algorithme de reconnaissance des chutes et des comportements à risque pouvant entraîner une chute, comme monter sur une chaise ou s'accroupir. Pour prévenir la fragilité, nous avons développé un algorithme de détection d'activité, pour savoir si la personne continue de réaliser autant d'activité qu'à son habitude, et un algorithme d'analyse de la marche, pour évaluer l'évolution de ses comportements locomoteurs. Ces deux algorithmes permettent d'établir un degré de fragilité des personnes, comme les médecins le font en consultation avec leurs tests cliniques. En résumé, nous avons contribué à la conception d'un système d'aide au maintien à domicile des personnes âgées, par le développement d'un algorithme de détection d'activité (capable, entre autres, de nous détecter les chutes et les situations à risque) et le développement d'un algorithme d'analyse de la marche.

Le système que nous avons imaginé doit, pour pouvoir être installé au domicile, être peu coûteux et ne pas perturber la vie quotidienne des personnes. Nous avons donc développé des algorithmes possibles à implémenter dans n'importe quelle machine bas coût et fonctionnant en temps réel pour pouvoir détecter les chutes le plus rapidement possible. Le système repose sur un capteur bas coût, non embarqué sur la personne. Nous avons choisi d'utiliser la caméra de profondeur de Microsoft. A partir de cette caméra, nous pouvons détecter la personne à son domicile avec un algorithme de soustraction du fond très simple et peu coûteux en temps de calcul. Le fond est appris à partir de l'image de profondeur et la personne est extraite de la scène par soustraction. Suivre la personne dans le temps peut requérir beaucoup de ressource, car chaque partie du corps a des mouvements indépendants qu'il est possible de suivre séparément. Par souci d'économie, nous extrayons quelques paramètres de l'image nous permettant de représenter la personne au cours du temps. Tout d'abord, nous nous intéressons au centre de masse, car certains auteurs ont mis en évidence que pour suivre le déplacement du corps en mouvement, la trajectoire du centre de masse est une information riche. A partir du centre de masse, nous pouvons obtenir des informations comme la position, la vitesse verticale et horizontale. Nous avons également

extrait deux autres paramètres qui ont une importance pour détecter l'activité, il s'agit de la dispersion verticale de la silhouette et du point maximum du corps. Ces 5 paramètres, simples à extraire, sont les caractéristiques représentant la personne au cours du temps et sur lesquels nos algorithmes de détection d'activité et d'analyse de la marche sont fondés.

Selon la littérature, l'analyse des paramètres de la marche, tels que les longueurs de pas, la cadence et la vitesse de marche sont de bons indicateurs de l'évolution de la fragilité des personnes âgées. L'algorithme d'extraction des paramètres de la marche d'une personne est fondé sur une seule caractéristique représentant la personne, à savoir le centre de masse. A partir du centre de masse et de la détection des maxima locaux sur l'axe vertical, nous obtenons les paramètres de la marche. En comparant les paramètres de la marche fournis par notre algorithme et ceux obtenus avec une méthode de référence, pour 10 sujets participant à l'expérience, nous obtenons des résultats assez précis puisque l'erreur obtenue est comparable à l'erreur due à la précision de la caméra.

L'algorithme de détection d'activité, fondé sur des modèles de Markov cachés, permet de discriminer entre 8 et 9 activités selon le modèle construit. Il permet de reconnaître les activités s'asseoir, marcher, s'accroupir, se pencher, monter sur un obstacle, s'allonger sur un lit ou un canapé, chuter, être allongé au sol et certains modèles identifient aussi l'activité être debout. Les modèles construits distinguent les activités à partir de 3 ou 5 observations représentant les personnes. Les modèles reconnaissant 8 activités se basent sur la position et vitesse du centre de masse et la distribution verticale de la silhouette. Les modèles reconnaissant 9 états se basent sur 5 paramètres en ajoutant la vitesse horizontale et le point maximum du corps. Une base de données a été construite en demandant à 28 sujets de réaliser des situations correspondantes aux activités devant être reconnues. Les paramètres des différents modèles construits, comme les probabilités de transition, les probabilités initiales, les fonctions d'observations sont apprises à partir de quelques séquences de la base de données. Les séquences restantes de la base servent à vérifier la validité des modèles pour classifier correctement les activités. Ainsi, nous avons pu constater que les modèles discriminant 8 activités classifient correctement toutes les situations sans faire de fausses détections, notamment sans déclarer une chute alors qu'il n'y en a pas. Les modèles détectant 9 activités ont des bons taux de classifications également mais indiquent trop souvent l'activité chute alors qu'elle n'est pas réalisée. Malgré l'augmentation du nombre d'observations pour discriminer les activités (5 pour les modèles discriminant 9 états, et 3 pour ceux discriminant 8 activités), l'augmentation de la complexité du modèle ne permet pas d'être plus performant pour reconnaître les activités, au contraire cela ajoute du bruit. En revanche, les modèles discriminant les activités à partir de 5 observations ont l'avantage d'être plus robustes aux situations où la personne n'est pas entièrement visible. Ces situations sont considérées comme nouvelles, car la base de données ne contenait que des situations où la personne était entièrement visible.

Les algorithmes de détection d'activité et d'analyse de la marche ont été testés sur quelques personnes âgées. Les tests montrent que certaines situations sont reconnues. Pour la suite, il est toutefois nécessaire de refaire des expérimentations avec des personnes âgées pour créer une nouvelle base de données adaptée à la bonne catégorie de personnes. De plus, même si les personnes âgées ont une façon de marcher différente des personnes jeunes, l'algorithme est toujours capable d'extraire les paramètres de la marche. Une étude plus approfondie permettra de déterminer si la précision des paramètres de la marche, extrait avec notre système, est similaire à la précision obtenue dans l'expérimentation avec les sujets jeunes. Ces premiers tests nous ont également permis de constater d'autres applications possibles du système. Notamment, l'observation de la trajectoire générale du centre de masse pourrait être un indicateur d'observance pour le médecin. La trajectoire pourrait également servir comme indicateur pour discriminer plusieurs

personnes dans une scène, car chaque trajectoire présente des irrégularités qui sont propres à chaque personne.

L'une des difficultés est de concevoir un système assez générique pour pouvoir être utilisé pour une population de personnes très disparate. Nous ne pouvons pas insérer manuellement des seuils à partir desquelles alerter les personnes d'une perte d'autonomie. Le système doit être capable de s'adapter à chaque personne, à partir d'une quantité suffisante de données, pour qu'il puisse apprendre ses habitudes au fur et à mesure. Il doit également apprendre les paramètres de la marche habituels et pouvoir ainsi détecter les seuils au-delà desquels l'évolution doit être considérée comme anormale. Ces seuils doivent être appris automatiquement par l'algorithme.

L'algorithme d'analyse de la marche et celui de détection d'activité peuvent être regroupés dans un système unique. Chaque caméra placée dans une pièce de la maison analyserait l'activité de la personne. Chaque détection des activités « Chute », « Accroupi » ou « Monté Obstacle » serait signalée aux personnes extérieures pour intervenir le plus rapidement possible en cas de chute. Les détections de l'activité « Marche » permettraient de lancer l'algorithme d'extraction des paramètres et de calculer le temps passé dans cette activité. Ce temps serait alors comparé au temps passé dans les situations « Assis » et « Allongé ». L'analyse de l'évolution des paramètres de la marche et l'évolution des activités constituerait des informations dans le cadre de la prévention de la fragilité.

Ce système a été pensé en vue d'être expérimenté au domicile des personnes âgées. Dans les objectifs, nous avons mentionné que nous voulions construire un système réaliste, utilisable. Se pose également la question de l'acceptation du système par les personnes. Peu d'études évaluent la perception qu'ont les personnes âgées face aux nouvelles technologies pouvant être installées à leur domicile. La question de l'acceptabilité est davantage nécessaire pour l'installation de caméras qui pose le problème du respect de la vie privée de la personne. Bellotti et Sellen [Bellotti and Sellen, 1993] précise qu'il est nécessaire de répondre à quatre catégories de questions lorsque l'on crée un système ayant pour objectif d'être installé au domicile. La réponse à ces problématiques peuvent influencer la façon dont la personne perçoit l'intrusion dans sa vie privée de la technologie, à savoir : quand et quelle information est enregistrée, comment l'information est enregistrée et une fois enregistrée ce qui lui arrive, qui a accès à l'information et comment l'information est utilisée. Les systèmes doivent être techniquement fiables et inspirer la confiance chez les utilisateurs. Malgré le peu d'articles existants sur l'acceptation des caméras à domicile, nous pouvons citer quelques auteurs proposant des débuts de réponse et des pistes à prendre en compte pour développer un système qui sera mieux accepté. Selon Caine *et al.* [Caine *et al.*, 2005], les personnes seraient prêtes à faire des compromis par rapport à leur vie privée si cela leur permet de rester indépendantes. Melenhorst *et al.* [Melenhorst *et al.*, 2004] ont également montré que les technologies perçues comme envahissantes pour l'intimité sont susceptibles d'être acceptées si elles sont considérées comme nécessaires pour répondre à un besoin. D'autres articles montrent que l'anonymat des images pourrait être une solution pour rendre les systèmes à base de caméra moins intrusifs dans l'intimité des personnes. Notamment l'étude de Demiris *et al.* [Demiris *et al.*, 2009] consistait à demander à 10 personnes d'effectuer plusieurs scénarios représentant des activités de la vie quotidienne comme s'asseoir dans le salon, accueillir un visiteur, préparer un repas, etc. Ces séquences étaient enregistrées par deux webcams et analysées de façon à rendre les images anonymes. Pour cela, ils ont utilisé une méthode de soustraction du fond pour obtenir uniquement la silhouette. Puis ils ont créé une deuxième représentation en modifiant la silhouette pour la transformer en 3 dimensions. Les deux représentations étaient alors montrées aux personnes lors d'interviews. Les interviews ont été menées autour des quatre problématiques qui ont été décrites par Bellotti et Sellen. Les réponses aux interviews ont montré que les personnes s'intéressaient à comprendre comment le système fonctionne, comment est

extraite la silhouette. Les interviews ont également indiqué que l'utilisation des vidéos réduite au simple suivi de la silhouette atténuait les préoccupations concernant le respect de la vie privée, car pour les personnes la vidéo restait anonyme. Par contre la représentation 3D n'a pas été appréciée. Selon les commentaires des personnes, cette représentation enlevait ce qui caractérise la personne, ils ne se reconnaissaient plus et donc ils ne voyaient pas comment le système pouvait être utilisé à des fins de sécurisation. Cela confirme la nécessité pour les utilisateurs d'avoir une compréhension non seulement de la sortie du système, mais aussi de ses processus et de son but. Elles ont également exprimé le désir de contrôler le fonctionnement du système en étant en mesure de l'éteindre et de l'allumer quand elles le désiraient. De plus, elles ont souligné l'importance de pouvoir déterminer elles-mêmes les personnes qui pourront avoir accès à l'information recueillie. Selon les personnes interrogées, certaines voulaient que les proches puissent avoir accès à l'information pour les rassurer et d'autres ne le souhaitaient pas. A la question de l'intérêt d'un tel système, neuf personnes sur dix ont répondu qu'elles ne voulaient pas installer ce système chez elles, car elles ne jugeaient pas en avoir besoin, elles se sentaient indépendantes. La seule personne qui voulait installer ce dispositif avait récemment chuté et a précisé qu'un tel système aurait été utile lors de sa chute pour alerter des personnes.

Pour construire le système final, nous devons donc prendre en compte ces différentes préconisations pour que le système soit accepté. Les images doivent rester anonymes. Pour cela, le système est conçu pour que l'analyse des images puisse être faite localement et que seules les mesures d'activité soient transmises ou stockées. Nous ajouterons au système la possibilité de l'éteindre ou l'allumer. Cette action est non seulement utile pour la personne suivie mais aussi pour les personnes extérieures (personnel soignant et famille) qui pourraient avoir des réticences à être filmées. L'utilisateur doit être mis dans la boucle en lui expliquant l'intérêt qu'il a d'installer un tel système, lui expliquer le fonctionnement et lui laisser la décision de qui aura accès aux données.

Annexe A

Modèle de Markov caché

Le modèle de Markov caché est la méthode paramétrique choisie pour développer un algorithme de détection d'activité. Avant de définir le formalisme, nous présentons le cas plus général des processus de Markov. Nous présentons également les différents algorithmes, de la littérature, utilisés pour l'apprentissage et l'inférence dans les modèles de Markov cachés.

A.1 Définition des modèles de Markov cachés

A.1.1 Processus de Markov

Un modèle de Markov est un processus stochastique dont le but est de représenter l'évolution dans le temps de l'état d'un système. Les modèles de Markov suivent l'hypothèse de Markov, selon laquelle l'état courant ne dépend que d'un nombre fini et fixé d'états précédents. En d'autres termes, l'état dans lequel se trouve le système au temps t ne dépend que de l'état dans lequel il se trouvait aux temps $t-1, t-2, \dots, t-n$, n étant le nombre d'états antérieurs que l'on souhaite prendre en compte. Les processus satisfaisant cette hypothèse sont appelés processus de Markov ou chaînes de Markov.

Deux éléments composent un modèle de Markov, les états et les transitions. Une transition est définie comme la probabilité de passer d'un état à un autre. A chaque pas de temps, une transition est appliquée au modèle qui modifie son état selon une certaine probabilité. La représentation d'un modèle de Markov, à deux états, est présentée à la figure A.1. Sur cette figure sont représentés deux états et les probabilités de transition entre chaque état. Par exemple, la probabilité de passer de l'état 1 à l'état 2 est de 70 %.

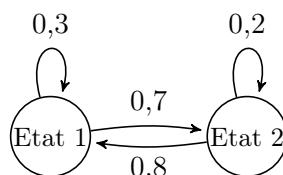


FIGURE A.1 – Exemple de modèle de Markov à deux états.

Les modèles de Markov permettent d'étudier l'évolution de la séquence d'états produite par le modèle, et surtout sa probabilité d'apparition. C'est pourquoi, une autre façon de représenter les chaînes de Markov est fondée sur le temps. La chaîne de Markov à la figure A.2 représente une séquence d'états de 0 à t , où X_i représente l'état à l'instant i .

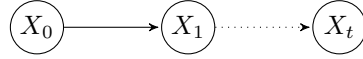


FIGURE A.2 – Représentation graphique d'une chaîne de Markov.

A.1.2 Modèle de Markov caché

Un modèle de Markov caché (MMC), en anglais *Hidden Markov Model* (HMM), est basé sur un modèle de Markov avec la particularité que les états ne sont pas directement observables [Rabiner, 1989]. Chaque état conditionne les observations qui sont visibles et qui permettront d'estimer l'état courant. Un MMC étudie l'évolution d'une séquence d'observations mesurées dans un système, comme représenté à la figure A.3. La figure illustre un MMC évoluant dans le temps avec X_t les états cachés, représentés en gris, et O_t les observations. Un MMC est défini par :

- un ensemble d'états;
- une matrice $a_{ij} = P(X_t = j | X_{t-1} = i)$ représentant toutes les probabilités de transition entre chaque paire d'états;
- une probabilité initiale $\pi_i = P(X_0 = i)$ pour chaque état i , qui est la probabilité *a priori* que le système se trouve dans l'état i à $t = 0$;
- une fonction d'observation $b_i(\mathbf{o}) = P(O_t = \mathbf{o} | X_t = i)$.

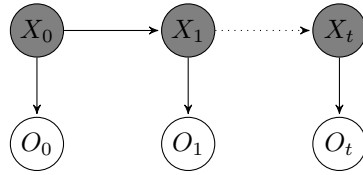


FIGURE A.3 – Représentation graphique d'une chaîne de Markov cachée, avec X_t le processus Markovien caché et O_t la séquence d'observations correspondantes.

A.2 Algorithmes d'apprentissage et d'inférence

L'objectif de cette section est de présenter les algorithmes capables d'inférer, pour un MMC, dans quel état se trouve la personne à chaque pas de temps. Les MMC peuvent être utilisés soit individuellement, soit par combinaison. Pour résoudre les systèmes basés sur un unique MMC, les algorithmes d'inférences couramment utilisés sont les algorithmes *Forward-Backward* et de Viterbi [Rabiner, 1989]. Pour résoudre les systèmes à base de combinaison de MMC, un algorithme d'inférence et un algorithme d'apprentissage, tel que l'algorithme de Baum-Welch, sont nécessaires.

A.2.1 Inférence pour le cas du MMC unique

Deux algorithmes peuvent résoudre l'inférence, l'algorithme *Forward-Backward* et l'algorithme de Viterbi.

Algorithme *Forward-Backward*

L'algorithme *Forward-Backward* est utilisé ici pour calculer la probabilité d'être dans tel ou tel état, connaissant les observations et le modèle. Le modèle constitue un ensemble d'hypothèses

que l'on définit comme vrai.

La probabilité de chaque état est représentée par la variable γ et l'équation qui en découle est la suivante :

$$\gamma_t(i) = P(q_t = i | O, \lambda)$$

avec q_t l'état au temps t , O la séquence d'observation $O = (o_1 o_2 \dots o_T)$ et λ le modèle. La variable γ calcule donc la probabilité d'être dans l'état i au temps t , sachant la séquence d'observation O et le modèle λ . Pour calculer la variable γ , il est nécessaire de calculer des variables intermédiaires obtenues grâce à deux étapes, *Forward* et *Backward*.

La phase *Forward* permet de calculer la probabilité jointe d'observer une séquence de valeurs passées et d'être dans tel état. La phase *Backward* calcule la probabilité d'être dans tel état sachant les observations futures. Ainsi, la probabilité γ de l'état sachant les observations passées et futures est calculée à partir des deux variables obtenues avec *Forward* et *Backward*.

Forward L'idée est de calculer la probabilité jointe de la séquence d'observation $O = (o_1, o_2, \dots, o_t)$ (jusqu'au temps t) et de l'état i au temps t , sachant le modèle λ . L'équation à résoudre est la suivante :

$$\alpha_t(i) = P(o_1, o_2, \dots, o_t, q_t = i | \lambda).$$

La première étape de l'algorithme initialise la probabilité α comme la probabilité jointe de l'état i et de l'observation initiale o_1 :

$$\alpha_1(i) = \pi_i b_i(o_1), \quad \text{pour } 1 \leq i \leq N$$

c'est-à-dire que la valeur α de l'état i au temps $t = 1$ est égale à π_i , la probabilité initiale de l'état i , multipliée par $b_i(o_1)$, la probabilité d'observation de l'état i .

Pour l'induction allant de $t > 1$ à $t = T$, le α est égal à :

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^N \alpha_t(i) a_{ij} \right] b_j(o_{t+1}), \quad \text{pour } 1 \leq t \leq T-1$$

et pour $1 \leq j \leq N$,

avec N le nombre d'état, a_{ij} la probabilité de transition entre i et j , $b_j(o_{t+1})$ la probabilité d'observation de l'état j au temps $t+1$. Ce calcul est représenté graphiquement à la figure A.4. Cette figure est une illustration, pour un MMC à trois états, du calcul de la probabilité effectué pour $\alpha_{t+1}(j = 1)$ par *Forward*.

Dans l'implémentation, une étape de normalisation est nécessaire pour permettre aux valeurs de rester dans un domaine représentable numériquement, soit :

$$\alpha'_{t+1}(j) = \frac{\alpha_{t+1}(j)}{\left[\sum_{j=1}^N \alpha_{t+1}(j) \right]}, \quad \text{pour } 1 \leq j \leq N.$$

Backward L'idée est de calculer la probabilité d'être dans un état sachant les observations futures. Contrairement à *Forward*, la probabilité à l'instant t est calculée à partir de la probabilité à l'instant $t+1$ c'est-à-dire que le calcul se fait à partir du dernier pas de temps T jusqu'au début de la séquence, sachant le modèle λ . L'équation à résoudre est :

$$\beta_t(i) = P(o_{t+1}, o_{t+2}, \dots, o_T, q_t = i | \lambda).$$

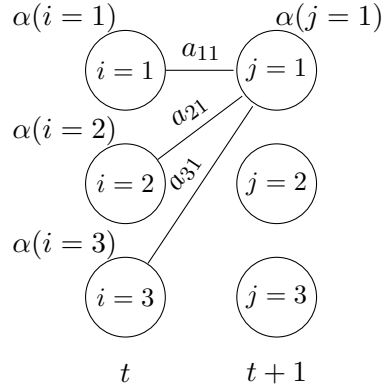


FIGURE A.4 – Représentation du calcul $\alpha_{t+1}(j = 1)$ par l'algorithme *Forward* pour un modèle à trois états.

A l'initialisation, le β vaut 1 pour tous les états :

$$\beta_T(i) = 1, \quad \text{pour } 1 \leq i \leq N.$$

Pour l'induction allant de $t = T - 1$ à $t = 1$, le β est égal à :

$$\beta_t(i) = \sum_{j=1}^N a_{ij} b_j(o_{t+1}) \beta_{t+1}(j), \quad \text{pour } t = T - 1, T - 2, \dots, 1$$

et pour $1 \leq i \leq N$,

avec N le nombre d'état, a_{ij} la probabilité de transition entre i et j , $b_j(o_{t+1})$ la probabilité d'observation de l'état j au temps $t + 1$. Ce calcul est représenté graphiquement à la figure A.5. Cette figure est une illustration, pour un MMC à trois états, du calcul de la probabilité effectuée pour $\beta_t(i = 1)$ par *Backward*.

Dans l'implémentation, une étape de normalisation est effectuée à chaque pas de temps soit :

$$\beta'_t(i) = \frac{\beta_t(i)}{\left[\sum_{i=1}^N \beta_t(i) \right]}, \quad \text{pour } 1 \leq i \leq N.$$

Couplage *Forward-Backward* Après avoir calculé α et β pour chaque pas de temps, γ est alors déduit à partir de ces deux valeurs pour ainsi donner les probabilités d'apparition de chaque état de $t = 1$ à $t = T$, sachant les observations passées et futures et le modèle. L'équation à calculer est la suivante :

$$\gamma_t(i) = \frac{\alpha_t(i) \beta_t(i)}{\sum_{i=1}^N \alpha_t(i) \beta_t(i)}, \quad \text{pour } 1 \leq i \leq N.$$

L'état le plus probable à chaque pas de temps est déduit en prenant la valeur du γ la plus grande. $\text{Max}(\gamma_t(i))$ conserve l'état parmi tous les états au temps t ayant le γ le plus élevé. Ainsi dans notre cas, nous obtiendrons l'état dans lequel la personne est passée. Et comme à chaque état du MMC correspond une activité, ainsi nous obtenons, en conservant l'état ayant la plus grande probabilité, l'activité qui est réalisée par la personne.

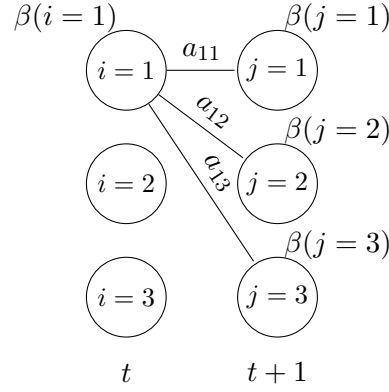


FIGURE A.5 – Représentation du calcul pour l'état $\beta_t(i = 1)$ par l'algorithme *Backward* pour un modèle à trois états.

Algorithme de Viterbi

L'algorithme de Viterbi, à la différence de l'algorithme *Forward-Backward*, fournit le « chemin » le plus probable, c'est-à-dire la meilleure séquence d'état $(q_1, q_2, \dots, q_T)$, sachant les observations et le modèle. Pour résoudre ce problème, il est nécessaire de passer par le calcul de δ défini par :

$$\delta_t(i) = \max_{q_1, q_2, \dots, q_{t-1}} P[q_1 q_2 \dots q_{t-1}, q_t = i, o_1 o_2 \dots o_t | \lambda].$$

A travers le $\delta_t(i)$, cet algorithme calcule pour chaque état caché du modèle, le chemin le plus probable atteignant l'état i au temps t . Pour mémoriser le chemin, l'état qui maximise les probabilités à chaque pas de temps sera conservé à travers la variable ψ .

Au temps $t = 1$, l'initialisation est définie par :

$$\begin{aligned} \delta_1(i) &= \pi_i b_i(o_1), & \text{pour } 1 \leq i \leq N \\ \psi_1(i) &= 0. \end{aligned}$$

Pour l'induction allant de $t > 1$ à $t = T$:

$$\begin{aligned} \delta_t(j) &= \max_{1 \leq i \leq N} [\delta_{t-1}(i) a(i, j)] b_j(o_t), & \text{pour } 2 \leq t \leq T, \\ \psi_t(j) &= \arg \max_{1 \leq i \leq N} [\delta_{t-1}(i) a(i, j)] & \text{et pour } 1 \leq j \leq N. \end{aligned}$$

$\psi_t(j)$ est donc la mémorisation de l'état précédent (au temps $t - 1$) appartenant au meilleur chemin qui mène à l'état j au temps t . La figure A.6 représente le calcul de δ et de ψ pour l'état j au temps t , effectué par l'algorithme de Viterbi pour un modèle à trois états. Le rouge symbolise l'état et la transition atteignant l'état j qui a la plus grande probabilité.

A l'implémentation, à chaque pas de temps et pour chaque état, il est nécessaire de normaliser :

$$\delta'_t(j) = \frac{\delta_t(j)}{\sum_{j=1}^N \delta_t(j)}, \quad \text{pour } 1 \leq j \leq N.$$

Au dernier pas de temps, pour $t = T$, la plus grande probabilité entre tous les états est calculée :

$$P^* = \max_{1 \leq i \leq N} [\delta_T(i)],$$

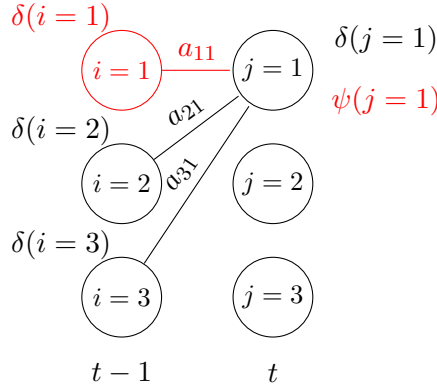


FIGURE A.6 – Représentation du calcul $\delta_t(j = 1)$ et $\psi_t(j = 1)$ par l’algorithme de Viterbi pour un modèle à trois états.

$$q_T^* = \arg \max_{1 \leq i \leq N} [\delta_T(i)].$$

Une dernière étape est nécessaire, il s’agit d’une étape de rétro-propagation. En prenant la plus haute probabilité parmi les états du temps $t = T$, on retrace le meilleur chemin en remontant jusqu’au premier pas de temps :

$$q_t^* = \psi_{t+1}(q_{t+1}^*), \quad \text{pour } t = T - 1, T - 2, \dots, 1.$$

La figure A.7 représente les différentes étapes effectuées par l’algorithme de Viterbi. A chaque étape de l’induction, l’algorithme se souvient de l’état le plus probable (rond rouge) et pour chaque état il se souvient de son prédécesseur le plus probable (transition rouge). Cette figure représente également la rétro-propagation à partir du dernier temps $t = T$ sur 4 pas de temps. Le chemin le plus probable trouvé par l’algorithme de Viterbi, sur cette figure, est représenté en vert. Les chiffres correspondent aux différentes itérations de l’étape d’induction (en rouge) et de l’étape de rétro-propagation (en vert). Nous pouvons remarquer que le chemin le plus probable ne correspond pas forcément aux transitions des états les plus probables (représentées en rouge) définies par les calculs de l’induction.

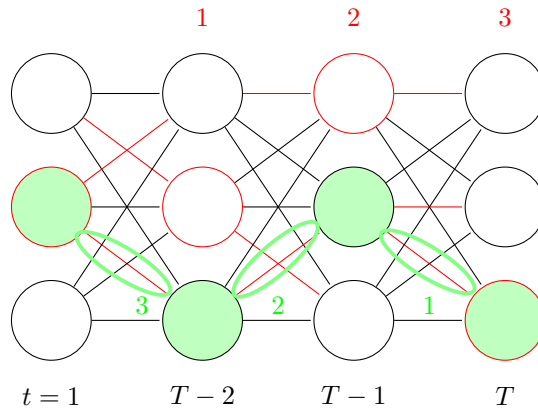


FIGURE A.7 – Représentation du calcul de l’algorithme de Viterbi avec la rétro-propagation (en vert) pour quatre pas de temps successifs et trois états.

A.2.2 Apprentissage et inférence pour les systèmes de plusieurs MMC

L'utilisation de MMC combiné permet de ne pas insérer de connaissance *a priori*. L'objectif est de générer un modèle appris en fonction des données qui lui sont fournies. La solution consiste à construire un modèle qui maximise la vraisemblance, c'est-à-dire la probabilité d'apparition d'une séquence d'observations $O = (o_1, o_2, \dots, o_T)$. En d'autres termes, l'objectif est de trouver le modèle qui explique le mieux la séquence.

Pour construire un modèle collant au plus près des données, il est nécessaire d'apprendre les paramètres du modèle qui sont les probabilités initiales, les probabilités de transition et les paramètres de la fonction d'observation (moyenne et covariance de chaque état). Pour estimer les paramètres, l'algorithme de Baum-Welch est utilisé. Cet algorithme présente le défaut de ne pas toujours converger vers la même solution.

Apprentissage

L'algorithme de Baum-Welch est un cas particulier de l'algorithme Espérance-Maximisation, qui sont les deux étapes à appliquer itérativement pour apprendre les paramètres du modèle. La première étape permet d'estimer la séquence d'états correspondant aux observations à partir du modèle courant. La seconde étape consiste à ré-estimer les paramètres du modèle pour augmenter sa vraisemblance. La base d'apprentissage est composée d'un ensemble d'exemples pour chaque état à modéliser.

Espérance Cette première étape consiste à calculer pour chaque séquence, les probabilités a posteriori. Cette étape est réalisée avec l'algorithme *Forward-Backward* décrit précédemment à la section A.2.1. Le résultat de cette étape est pour chaque séquence s de longueur T_s , composant les exemples de l'activité à modéliser, les valeurs α_t^s , β_t^s et γ_t^s .

Pour la première itération, les valeurs prises par le modèle sont des valeurs d'initialisation arbitraires. Pour les probabilités initiales de chaque état, nous les fixons à l'identique. Concernant la matrice de transition, toutes les probabilités sont fixées également de manière équivalente. Concernant les observations, la moyenne et la covariance, nécessaires pour former les gaussiennes multivariées de chaque état et chaque MMC, doivent être initialisées. Les moyennes pour la position du centre de masse, la vitesse, et la distribution verticale sont initialisées en prenant, de manière aléatoire, les valeurs d'une observation correspondant à une activité. A partir de ces moyennes tirées au hasard, la covariance est calculée pour chaque état.

Pour les itérations suivantes, l'algorithme consiste à reprendre les paramètres estimés par l'étape de maximisation.

Maximisation Dans cette étape, la moyenne et la covariance sont alors recalculées de façon à maximiser la vraisemblance, pour faire en sorte que le modèle soit plus proche de la réalité. En d'autres termes, l'ensemble des MMC est réévalué par rapport à la séquence d'observations sur laquelle nous voulons l'entraîner. L'estimation des paramètres est renouvelée jusqu'à ce qu'il y ait convergence.

Réestimation de la matrice de transition Les probabilités de transition sont recalculées en comptant les transitions à partir des données obtenues à l'étape Espérance. L'équation suivante permet de calculer $\xi_t(i, j)$, qui est la probabilité de transition entre les états i et j entre les instants

t et $t + 1$.

$$\xi_t^s(i, j) = \frac{\alpha_t^s(i) a_{ij} b_j(\mathbf{o}_{t+1}^s) \beta_{t+1}^s(j)}{\sum_{k=1}^N \sum_{l=1}^N \alpha_t^s(k) a_{kl} b_l(\mathbf{o}_{t+1}^s) \beta_{t+1}^s(l)}$$

où s est le numéro de la séquence dans la base d'apprentissage.

La probabilité de transition entre i et j est ensuite obtenue en faisant la somme sur toutes les séquences des $\xi_t^s(i, j)$ divisée par la somme des $\gamma_t^s(i)$ qui représente la probabilité de se trouver dans l'état i :

$$a_{ij} = \frac{\sum_{s=1}^S \sum_{t=1}^{T_s-1} \xi_t^s(i, j)}{\sum_{s=1}^S \sum_{t=1}^{T_s-1} \gamma_t^s(i)}.$$

Réestimation de la matrice des probabilités initiales La probabilité initiale est obtenue en regardant la probabilité des différents états en début de séquences dans les données de l'étape Espérance :

$$\pi_i = \frac{\sum_{s=1}^S \gamma_0^s(i)}{S}.$$

Réestimation de la matrice des probabilités d'observations. Les moyennes et covariances des fonctions d'observations sont recalculées pour chaque état en prenant en compte à chaque instant la probabilité d'appartenir à l'état :

$$\begin{aligned} \mu_i &= \frac{\sum_{s=1}^S \sum_{t=1}^{T_s} \gamma_t^s(i) \mathbf{o}_t^s}{\sum_{s=1}^S \sum_{t=1}^{T_s} \gamma_t^s(i)}, \\ \Sigma_i &= \frac{\sum_{s=1}^S \sum_{t=1}^{T_s} \gamma_t^s(i) (\mathbf{o}_t^s - \mu_i)(\mathbf{o}_t^s - \mu_i)'}{\sum_{s=1}^S \sum_{t=1}^{T_s} \gamma_t^s(i)}. \end{aligned}$$

Inférence

L'inférence est réalisée en calculant à chaque pas de temps la vraisemblance du modèle. L'état le plus probable est celui qui correspond au modèle donnant la plus grande vraisemblance. La vraisemblance L se calcule en reprenant les valeurs de l'étape *Forward* en fin de séquence :

$$L = \sum_{i=1}^N \alpha_T(i).$$

Dans l'implémentation, il est habituel d'utiliser le logarithme de la vraisemblance qui permet de garder les valeurs dans un domaine représentable numériquement. Et comme les valeurs α ont été normalisées à chaque pas de temps, le logarithme de la vraisemblance se calcule en faisant la somme des logarithmes des coefficients de normalisation.

Publications

- [1] Amandine Dubois and François Charpillet. A gait analysis method based on a depth camera for fall prevention. In *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2014.
- [2] Amandine Dubois and François Charpillet. Detecting and preventing falls with depth camera, tracking the body center. In *12th European Association for the Advancement of Assistive Technology in Europe (AAATE)*, 2013.
- [3] Amandine Dubois and François Charpillet. Human activities recognition with RGB-Depth camera using HMM. In *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4666–4669. IEEE, 2013.
- [4] Amandine Dubois and François Charpillet. Automatic fall detection system with a RGB-D camera using a Hidden Markov Model. In *11th International Conference on Smart Homes and Health Telematics (ICOST)*, pages 259–266, 2013.
- [5] Amandine Dubois and François Charpillet. Système d’évaluation de la fragilité chez les personnes âgées. In *Journée d’étude sur la Télésanté (JETSAN)*. IEEE france, 2013.
- [6] Amandine Dubois and François Charpillet. Tracking mobile objects with several kinects using HMMs and component labelling. In *Workshop Assistance and Service Robotics in a human environment, International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2012.
- [7] Amandine Dubois, Abdallah Dib, and François Charpillet. Using HMMs for discriminating mobile from static objects in a 3D occupancy grid. In *23rd International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 170–176. IEEE, 2011.

Références

- [Aggarwal and Ryoo, 2011] J.K. Aggarwal and Michael S. Ryoo. Human activity analysis : A review. *ACM Computing Surveys*, 43(3), 2011.
- [Alexander *et al.*, 1997] Neil B. Alexander, Jessica Ulbrich, Aarti Raheja, and Dwight Channer. Rising from the floor in older adults. *Journal of the American Geriatrics Society*, 45(5) :564–569, 1997.
- [Allard and Blanchi, 1996] Paul Allard and Jean-Pierre Blanchi. *Analyse du mouvement humain par la biomécanique*. Décarie, 1996.
- [Anderson *et al.*, 2006] Derek Anderson, James M. Keller, Marjorie Skubic, Xi Chen, and Zhi-hai He. Recognizing falls from silhouettes. In *28th Annual International Conference of the Engineering in Medicine and Biology Society, (EMBS)*, pages 6388–6391. IEEE, 2006.
- [Anderson *et al.*, 2009] Derek Anderson, Robert H. Luke, James M. Keller, Marjorie Skubic, Marilyn Rantz, and Myra Aud. Linguistic summarization of video for fall detection using voxel person and fuzzy logic. *Computer Vision and Image Understanding*, 113(1) :80–89, 2009.
- [André *et al.*, 2001] J.-M. André, J. Paysant, N. Martinet, J.-M. Beis, and C. Beyaert. Georges Gilles de la Tourette, initiateur de l’enregistrement de la marche dans les maladies du système nerveux. *Revue neurologique*, 157(3) :293–296, 2001.
- [Auvinet *et al.*, 2003] Bernard Auvinet, Gilles Berrut, Claude Touzard, Nadine Collet, Denis Chaleil, and Eric Barrey. Gait abnormalities in elderly fallers. *Journal of Aging and Physical Activity*, pages 40–52, 2003.
- [Auvinet *et al.*, 2011b] Edouard Auvinet, Franck Multon, Alain Saint-Arnaud, Jacqueline Rousseau, and Jean Meunier. Fall detection with multiple cameras : An occlusion-resistant method based on 3-D silhouette vertical distribution. *IEEE transactions on information technology in biomedicine : a publication of the IEEE EMBS*, 15(2) :290–300, 2011.
- [Avidan, 2004] Shai Avidan. Support vector tracking. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(8) :1064–1072, 2004.
- [Bellotti and Sellen, 1993] Victoria Bellotti and Abigail Sellen. Design for privacy in ubiquitous computing environments. In *Proceedings of the Third European Conference on Computer-Supported Cooperative Work (ECSCW)*, pages 77–92. Springer, 1993.
- [Bian *et al.*, 2012] Zhen-Peng Bian, Lap-Pui Chau, and Nadia Magnenat-Thalmann. Fall detection based on skeleton extraction. In *Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry*, pages 91–94. ACM, 2012.
- [Blanpain and Chardon, 2010] Nathalie Blanpain and Olivier Chardon. Projections de population à l’horizon 2060. Un tiers de la population âgé de plus de 60 ans, INSEE. Octobre 2010.

- [Bourke *et al.*, 2007] A.K. Bourke, J.V. O'Brien, and G.M. Lyons. Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. *Gait & posture*, 26(2) :194–199, 2007.
- [Bramell-Risberg *et al.*, 2005] Eva Bramell-Risberg, G-B Jarnlo, Lennart Minthon, and Sölve Elmståhl. Lower gait speed in older women with dementia compared with controls. *Dementia and geriatric cognitive disorders*, 20(5) :298–305, January 2005.
- [Caine *et al.*, 2005] Kelly E. Caine, Wendy A. Rogers, and Arthur D. Fisk. Privacy perceptions of an aware home with visual sensing devices. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 49, pages 1856–1858. SAGE Publications, 2005.
- [Camicioli *et al.*, 1998] R. Camicioli, D. Howieson, B. Oken, G. Sexton, and J. Kaye. Motor slowing precedes cognitive impairment in the oldest old. *American Academy of Neurology*, 50(5) :1496–1498, 1998.
- [Comaniciu, 2002] Dorin Comaniciu. Bayesian kernel tracking. In *Pattern Recognition*, pages 438–445. Springer, 2002.
- [Condouret *et al.*, 1987] J. Condouret, M. Iehl, C. F. Roques, P. Dupui, F. Montoya, B. Pages, P. Bessou, and M. Pujol. Analyse spatio-temporelle de la marche par la technique de Bessou : résultats chez l'hémiplégique. In *Annales de réadaptation et de médecine physique*, volume 30, pages 267–278. Elsevier, 1987.
- [Deeb *et al.*, 2012] Rada Deeb, Félix Lédée, Elodie Dessérée, and Saida Bouakaz. Méthode robuste pour la détection de chute dans un environnement non-contrôlé. In *Actes de la conférence RFIA, Reconnaissance des Formes et Intelligence Artificielle*, 2012.
- [Demiris *et al.*, 2009] George Demiris, Debra Parker Oliver, Jarod Giger, Marjorie Skubic, and Marilyn Rantz. Older adults' privacy considerations for vision based recognition methods of eldercare applications. *Technology and Health Care*, 17(1) :41–48, 2009.
- [Enzweiler and Gavrilu, 2009] Markus Enzweiler and Dariu M. Gavrilu. Monocular pedestrian detection : Survey and experiments. *Transactions on Pattern Analysis and Machine Intelligence*, 31(12) :2179–2195, 2009.
- [Faivre, 2003] Arnaud Faivre. *Conception et validation d'un nouvel outil d'analyse de la marche*. PhD thesis, Université de Franche-Comté, 2003.
- [Ferri *et al.*, 2006] Cleusa P. Ferri, Martin Prince, Carol Brayne, Henry Brodaty, Laura Fratiglioni, Mary Ganguli, Kathleen Hall, Kazuo Hasegawa, and Hugh Hendrie. Global prevalence of dementia : a Delphi consensus study. *The Lancet Neurology*, 2006.
- [Fried *et al.*, 2001] Linda P. Fried, Catherine M. Tangen, Jeremy Walson, Anne B. Newman, Calvin Hirsch, John Gottdiener, Teresa Seeman, Tracy Russell, Willem J. Kop, Gregory Burke, and Mary Ann McBurnie. Frailty in older adults : Evidence for a phenotype. *Journal of Gerontology Series A : Biological Sciences and Medical Sciences*, 56(3) :146–156, 2001.
- [Fusco, 2008] Nicolas Fusco. *Analyse, modélisation et simulation de la marche pathologique*. PhD thesis, Université Rennes 2, 2008.
- [Gaxatte *et al.*, 2011] C. Gaxatte, T. Nguyen, F. Chourabi, J. Salleron, V. Pardessus, I. Delabrière, A. Thévenon, and F. Puisieux. Fear of falling as seen in the multidisciplinary falls consultation. *Annals of physical and rehabilitation medicine*, 54(4) :248–258, 2011.
- [Gillet, 2004] Christophe Gillet. *Analyse Biomécanique de la marche et proposition de classes de marcheurs -Application au portage de sacs à dos*. PhD thesis, Université de Valenciennes et du Hainaut-Cambrésis, 2004.

- [Gómez-Conde, 2011] Iván Gómez-Conde. Simple human gesture detection and recognition using a feature vector and a real-time histogram based algorithm. *Journal of Signal and Information Processing*, 02(04) :279–286, 2011.
- [Grewe and Kak, 1995] Lynne Grewe and Avinash C. Kak. Interactive learning of a multiple-attribute hash table classifier for fast object recognition. *Computer Vision and Image Understanding*, 61(3) :387–416, 1995.
- [Guimaraes and Isaacs, 1980] R.M. Guimaraes and Bernard Isaacs. Characteristics of the gait in old people who fall. *Disability & Rehabilitation*, 2(4) :177–180, 1980.
- [Hageman and Blanke, 1986] Patricia A. Hageman and Daniel J. Blanke. Comparison of gait of young women and elderly women. *Physical therapy*, 66(9) :1382–1387, September 1986.
- [Hagler *et al.*, 2010] Stuart Hagler, Daniel Austin, Tamara L. Hayes, Jeffrey Kaye, and Misha Pavel. Unobtrusive and ubiquitous in-home monitoring : A methodology for continuous assessment of gait velocity in elders. *IEEE transactions on bio-medical engineering*, 57(4) :813–820, April 2010.
- [Harris and Stephens, 1988] Chris Harris and Mike Stephens. A combined corner and edge detector. In *4th Alvey Vision Conference*, volume 15, pages 147–151. Manchester, UK, 1988.
- [Hausdorff *et al.*, 2001] Jeffrey M. Hausdorff, Dean A. Rios, and Helen K. Edelberg. Gait variability and fall risk in community-living older adults : A 1-year prospective study. *Archives of physical medicine and rehabilitation*, 82(8) :1050–1056, August 2001.
- [Hewson *et al.*, 2007] David J. Hewson, Jacques Duchêne, François Charpillat, Jamal Saboune, Valérie Michel-Pellegrino, Hassan Amoud, Michel Doussot, Jean Paysant, Anne Boyer, and Jean-Yves Hogrel. The PARACHUTE project : remote monitoring of posture and gait for fall prevention. *EURASIP Journal on Advances in Signal Processing*, 2007.
- [Holtzer *et al.*, 2006] Roe Holtzer, Joe Verghese, Xiaonan Xue, and Richard B. Lipton. Cognitive processes related to gait velocity : Results from the einstein aging study. *Neuropsychology*, 20(2) :215–223, March 2006.
- [Jansen *et al.*, 2007] Bart Jansen, Frederik Temmermans, and Rudi Deklerck. 3D human pose recognition for home monitoring of elderly. In *29th Annual International Conference of the Engineering in Medicine and Biology Society (EMBS)*, pages 4049–4051. IEEE, January 2007.
- [Jepson *et al.*, 2003] Allan D. Jepson, David J. Fleet, and Thomas F. El-Maraghi. Robust online appearance models for visual tracking. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(10) :1296–1311, 2003.
- [Ke *et al.*, 2013] Shian-Ru Ke, Hoang Le Uyen Thuc, Yong-Jin Lee, Jenq-Neng Hwang, Jang-Hee Yoo, and Kyoung-Ho Choi. A review on video-based human activity recognition. *Computers*, 2(2) :88–131, 2013.
- [Kirtley, 2006] Christopher Kirtley. *Clinical Gait Analysis : Theory and practice*. Elsevier Health Sciences, 2006.
- [Kuo *et al.*, 2007] Hsu-Ko. Kuo, Suzanne G. Leveille, Yau-Hua Yu, and William P. Milberg. Cognitive function, habitual gait speed, and late-life disability in the national health and nutrition examination survey (NHANES) 1999-2002. *Gerontology*, 53(15) :102–110, 2007.
- [Lowe, 2004] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2) :91–110, 2004.
- [Marey, 1894] Etienne-Jules Marey. *Le mouvement*. Masson, 1894.

- [Martin *et al.*, 2004] David R. Martin, Charless C. Fowlkes, and Jitendra Malik. Learning to detect natural image boundaries using local brightness, color, and texture cues. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(5) :530–549, 2004.
- [Melenhorst *et al.*, 2004] Anne-Sophie Melenhorst, Arthur D. Fisk, Elizabeth D. Mynatt, and Wendy A. Rogers. Potential intrusiveness of aware home technology : Perceptions of older adults. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 48, pages 266–270. SAGE Publications, 2004.
- [Moravec, 1979] Hans P. Moravec. Visual mapping by a robot rover. In *Proceedings of the 6th International Joint Conference on Artificial Intelligence-Volume 1*, pages 598–600. Morgan Kaufmann Publishers Inc., 1979.
- [Nait-Charif and McKenna, 2004] H. Nait-Charif and S.J. McKenna. Activity summarisation and fall detection in a supportive home environment. In *17th International Conference on Pattern Recognition (ICPR)*, volume 4, pages 323–326. IEEE, 2004.
- [Noury *et al.*, 2007] Norbert Noury, Anthony Fleury, Pierre Rumeau, A.K. Bourke, G.O. Laighin, Vincent Rialle, and J.E. Lundy. Fall detection-principles and methods. In *Engineering in Medicine and Biology Society (EMBS), 29th Annual International Conference of the IEEE*, pages 1663–1666. IEEE, 2007.
- [Papageorgiou *et al.*, 1998] Constantine P. Papageorgiou, Michael Oren, and Tomaso Poggio. A general framework for object detection. In *International conference on Computer Vision, 1998*, pages 555–562. IEEE, 1998.
- [Pepin *et al.*, 2009] Nicolas Pepin, Olivier Simonin, and François Charpillet. Intelligent tiles : Putting situated multi-agents models in real world. In *International Conference on Agents and Artificial Intelligence (ICAART’09)*, pages 513–519, 2009.
- [Perry and Davids, 1992] Jacquelin Perry and Jon R. Davids. Gait analysis : Normal and pathological function. *Journal of Pediatric Orthopaedics*, 12(6) :815, 1992.
- [Peursum *et al.*, 2005] Patrick Peursum, Hung H. Bui, Svetha Venkatesh, and Geoff West. Robust recognition and segmentation of human actions using HMMs with missing observations. *EURASIP Journal on Applied Signal Processing*, 2005(13) :2110–2126, 2005.
- [Rabiner, 1989] Lawrence Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. In *Proceedings of the IEEE*, pages 257–286, 1989.
- [Rangarajan and Shah, 1991] Krishnan Rangarajan and Mubarak Shah. Establishing motion correspondence. *Conference Vision Graphics Image Process*, 54(1) :56–73, 1991.
- [Rimmer *et al.*, 2005] E. Rimmer, M. Wojciechowska, C. Stave, A. Sganga, and B. O’Connell. Implications of the facing dementia survey for the general population, patients and caregivers across europe. *International Journal of Clinical Practice*, 1(146) :17–24, March 2005.
- [Rimminen *et al.*, 2009] Henry Rimminen, Juha Lindström, and Raimo Sepponen. Positioning accuracy and multi-target separation with a human tracking system using near field imaging. *International Journal on Smart Sensing and Intelligent System*, 2(1) :156–175, 2009.
- [Rimminen *et al.*, 2010] Henry Rimminen, Juha Lindström, Matti Linnavuo, and Raimo Sepponen. Detection of falls among the elderly by a floor sensor using the electric near field. *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society*, 14(6) :1475–1476, November 2010.
- [Rose *et al.*, 2008] Cédric Rose, Jamal Saboune, and François Charpillet. Reducing particle filtering complexity for 3D motion capture using dynamic bayesian networks. In *Twenty-Third AAAI Conference on Artificial Intelligence*, pages 1396–1401, 2008.

- [Rougier and Meunier, 2010] Caroline Rougier and Jean Meunier. 3D head trajectory using a single camera. In *Image and Signal Processing*, pages 505–512. Springer, 2010.
- [Rougier *et al.*, 2011] Caroline Rougier, Edouard Auvinet, Jacqueline Rousseau, Max Mignotte, and Jean Meunier. Fall detection from depth map video sequences. *International Conference on Smart Homes and Health Telematics (ICOST)*, pages 121–128, 2011.
- [Rowley *et al.*, 1998] Henry A. Rowley, Shumeet Baluja, and Takeo Kanade. Neural network-based face detection. *Pattern Analysis and Machine Intelligence*, 20(1) :23–38, 1998.
- [Russell and Norvig, 2010] Stuart Jonathan Russell and Peter Norvig. *Intelligence artificielle*. Pearson Education France, 2010.
- [Saboune and Charpillet, 2005] Jamal Saboune and Francois Charpillet. Using interval particle filtering for marker less 3D human motion capture. In *17th IEEE International Conference on Tools with Artificial Intelligence (ICTAI’05)*. IEEE, 2005.
- [Saunders *et al.*, 1953] M. Saunders, Verne T. Inman, and Howard D. Eberhart. The major determinants in normal and pathological gait. *The Journal of Bone & Joint Surgery*, pages 543–558, July 1953.
- [Sethi and Jain, 1987] Ishwar K. Sethi and Ramesh Jain. Finding trajectories of feature points in a monocular image sequence. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (1) :56–73, 1987.
- [Shaw, 2002] Fiona E. Shaw. Falls in cognitive impairment and dementia. *Clinics Geriatric Medicine*, pages 18 :159–280, 2002.
- [Shi and Malik, 2000] Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. *Pattern Analysis and Machine Intelligence*, 22(8) :888–905, 2000.
- [Shotton *et al.*, 2013] Jamie Shotton, Toby Sharp, Alex Kipman, Andrew Fitzgibbon, Mark Finocchio, Andrew Blake, Mat Cook, and Richard Moore. Real-time human pose recognition in parts from single depth images. *Communications of the ACM*, 56(1) :116–124, 2013.
- [Silva *et al.*, 2002] Mauricio Silva, Eric F. Shepherd, Walter O. Jackson, Frederick J. Dorey, and Thomas P. Schmalzried. Average patient walking activity approaches 2 million cycles per year : pedometers under-record walking activity. *The Journal of Arthroplasty*, 17(6) :693–697, 2002.
- [Stauffer and Grimson, 2000] Chris Stauffer and W. Eric L. Grimson. Learning patterns of activity using real-time tracking. *Pattern Analysis and Machine Intelligence*, 22(8) :747–757, 2000.
- [Studenski *et al.*, 2003] Stephanie Studenski, Subashan Perera, Dennis Wallace, Julie M. Chandler, Pamela W. Duncan, Earl Rooney, Michael Fox, and Jack M. Guralnik. Physical performance measures in the clinical setting. *Journal of the American Geriatrics Society*, 51(3) :314–322, March 2003.
- [Suzuki *et al.*, 2003] Kenji Suzuki, Isao Horiba, and Noboru Sugie. Linear-time connected-component labeling based on sequential local operations. *Computer Vision and Image Understanding*, 89(1) :1–23, January 2003.
- [Thomesse *et al.*, 2001] Jean-Pierre Thomesse, David Bellot, Anne Boyer, Eric Campo, Marie Chan, François Charpillet, Jocelyne Fayn, Claire Leschi, Norbert Noury, Vincent Rialle, Laurent Romary, Paul Rubel, Nazha Selmaoui, François Steenkeste, and Gilles Virone. Integrated information technologies for patients remote follow-up and homecare. *3rd International Workshop on Enterprise Networking and Computing in Health Care Industry (HealthCom)*, 2001.

- [Tieu and Viola, 2004] Kinh Tieu and Paul Viola. Boosting image retrieval. *International Journal of Computer Vision*, 56(1-2) :17–36, 2004.
- [Tinetti *et al.*, 1993] Mary E. Tinetti, Wen-Liang Liu, and Elizabeth B. Claus. Predictors and prognosis of inability to get up after falls among elderly persons. *Journal of the American Medical Association (Jama)*, 269(1) :65–70, 1993.
- [Turaga *et al.*, 2008] Pavan Turaga, Rama Chellappa, Venkatramana S. Subrahmanian, and Octavian Udrea. Machine recognition of human activities : A survey. *IEEE Transactions on Circuits and Systems for Video Technology*, 18(11) :1473–1488, 2008.
- [Valera and Velastin, 2005] M. Valera and Sergio A. Velastin. Intelligent distributed surveillance systems : a review. In *Vision, Image and Signal Processing, IEE Proceedings-*, volume 152, pages 192–204. IET, 2005.
- [Van Iersel *et al.*, 2004] M.B. Van Iersel, W. Hoefsloot, M. Munneke, B.R. Bloem, and M.G.M. Olde Rikkert. Systematic review of quantitative clinical gait analysis in patients with dementia. *Zeitschrift für Gerontologie und Geriatrie*, 37(1) :27–32, February 2004.
- [Veenman *et al.*, 2001] Cor J. Veenman, Marcel J.T. Reinders, and Eric Backer. Resolving motion correspondence for densely moving points. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(1) :54–72, 2001.
- [Viel, 2000] Eric Viel. *La marche humaine, la course et le saut. Biomécanique, explorations, normes et dysfonctionnements*. Elsevier Masson, Paris, 2000.
- [Viola *et al.*, 2003] Paul Viola, Michael J. Jones, and Daniel Snow. Detecting pedestrians using patterns of motion and appearance. In *Proceedings of the International Conference on Computer Vision*, pages 734–741. IEEE, 2003.
- [Waite *et al.*, 2005] L.M. Waite, D.A Grayson, O. Piguet, H. Creasey, H.P. Bennett, and G.A. Broe. Gait slowing as a predictor of incident dementia : 6-year longitudinal data from the sydney older persons study. *Journal of the Neurological Sciences*, 229-230 :89–93, March 2005.
- [Webe, 2006] Amandine Webe. *Dépendance des personnes âgées et handicap : les opinions des Français entre 2000 et 2005*. DREES Études et résultats, mai 2006.
- [Wild *et al.*, 1981] D. Wild, U.S. Nayak, and B. Isaacs. How dangerous are falls in old people at home? *British Medical Journal (Clinical research ed.)*, 282(6260) :266, 1981.
- [Winter *et al.*, 1990] David A. Winter, Aftab E. Patla, James S. Frank, and Sharon E. Walt. Biomechanical walking pattern changes in the fit and healthy elderly. *Physical therapy*, 70(6) :340–347, June 1990.
- [Wu, 2000] Ge Wu. Distinguishing fall activities from normal activities by velocity characteristics. *Journal of Biomechanics*, 33(11) :1497–1500, 2000.
- [Yilmaz *et al.*, 2006] Alper Yilmaz, Omar Javed, and Mubarak Shah. Object tracking : A survey. *ACM Computing Surveys*, 38(4) :13, dec 2006.

Résumé

Le vieillissement de la population est un enjeu majeur pour les prochaines années en raison, notamment, de l'augmentation du nombre de personnes dépendantes. La question du maintien à domicile de ces personnes se pose alors, du fait de l'impossibilité pour les instituts spécialisés de les accueillir toutes et, surtout, de la volonté des personnes âgées de rester chez elles le plus longtemps possible. Or, le développement de systèmes technologiques peut aider à résoudre certains problèmes comme celui de la sécurisation en détectant les chutes, et de l'évaluation du degré d'autonomie pour prévenir les accidents. Plus particulièrement, nous nous intéressons au développement des systèmes ambiants, peu coûteux, pour l'équipement du domicile. Les caméras de profondeur permettent d'analyser en temps réel les déplacements de la personne. Nous montrons dans cette thèse qu'il est possible de reconnaître l'activité de la personne et de mesurer des paramètres de sa marche à partir de l'analyse de caractéristiques simples extraites des images de profondeur. La reconnaissance d'activité est réalisée à partir des modèles de Markov cachés, et permet en particulier de détecter les chutes et des activités à risque. Lorsque la personne marche, l'analyse de la trajectoire du centre de masse nous permet de mesurer les paramètres spatio-temporels pertinents pour l'évaluation de la fragilité de la personne. Ce travail a été réalisé sur la base d'expérimentations menées en laboratoire, d'une part, pour la construction des modèles par apprentissage automatique et, d'autre part, pour évaluer la validité des résultats. Les expérimentations ont montré que certains modèles de Markov cachés, développés pour ce travail, sont assez robustes pour classifier les différentes activités. Nous donnons, également dans cette thèse, la précision, obtenue avec notre système, des paramètres de la marche en comparaison avec un tapis actimétrique. Nous pensons qu'un tel système pourrait facilement être installé au domicile de personnes âgées, car il repose sur un traitement local des images. Il fournit, au quotidien, des informations sur l'analyse de l'activité et sur l'évolution des paramètres de la marche qui sont utiles pour sécuriser et évaluer le degré de fragilité de la personne.

Mots-clés: Caméra de profondeur, Modèle de Markov caché, Reconnaissance d'activité, Analyse de la marche.

Abstract

Population ageing is a major issue for society in the next years, especially because of the increase of dependent people. The limits in specialized institutes capacity and the wish of the elderly to stay at home as long as possible explain a growing need for new specific at home services. Technologies can help securing the person at home by detecting falls. They can also help in the evaluation of the frailty for preventing future accidents. This work concerns the development of low cost ambient systems for helping the stay at home of elderly. Depth cameras allow analysing in real time the displacement of the person. We show that it is possible to recognize the activity of the person and to measure gait parameters from the analysis of simple feature extracted from depth images. Activity recognition is based on Hidden Markov Models and allows detecting at risk behaviours and falls. When the person is walking, the analysis of the trajectory of her centre of mass allows measuring gait parameters that can be used for frailty evaluation. This work is based on laboratory experimentations for the acquisition of data used for models training and for the evaluation of the results. We show that some of the developed Hidden Markov Models are robust enough for classifying the activities. We also evaluate the precision of the gait parameters measurement in comparison to the measures provided by an actimetric carpet. We believe that such a system could be installed in the home of the elderly because it relies on a local processing of the depth images. It would be able to provide daily information on the person activity and on the evolution of her gait parameters that are useful for securing her and evaluating her frailty.

Keywords: Depth camera, Hidden Markov Model, Activity recognition, Gait analysis.

